



Suivi de pollution atmosphérique par système multi-capteurs - méthode mixte de classification et de détermination d'un indice de pollution..

Zaher Al Barakeh

► To cite this version:

Zaher Al Barakeh. Suivi de pollution atmosphérique par système multi-capteurs - méthode mixte de classification et de détermination d'un indice de pollution... Autre. Ecole Nationale Supérieure des Mines de Saint-Etienne, 2012. Français. NNT : 2012EMSE0677 . tel-00795609

HAL Id: tel-00795609

<https://theses.hal.science/tel-00795609>

Submitted on 28 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NNT : 2012 EMSE 0677

THÈSE

Présentée par

Zaher AL BARAKEH

Pour obtenir le grade de

Docteur de l'École Nationale Supérieure des Mines de Saint-Étienne

Spécialité : Génie Des Procédés

Suivi de pollution atmosphérique par système multi-capteurs – méthode mixte de classification et de détermination d'un indice de pollution.

Soutenue à Saint Etienne, le 17 décembre 2012

Membres du jury

Président :	Pierre LANTERI	Professeur, UCB-Lyon 1, Lyon
Rapporteurs :	Jacques NICOLAS	Professeur, ULG, Arlon
	Marcel BOUVET	Professeur, ICMUB, Dijon
Examineurs :	Laure CHEVALLIER	Docteur, Air Languedoc Roussillon, Montpellier
Directeurs de thèse :	Christophe PIJOLAT	Professeur, EMSE, Saint Etienne
	Nadine LOCOGE	Professeur, EMD, Douai
Encadrants de thèse :	Philippe BREUIL	Ingénieur de Recherche, EMSE, Saint Etienne
	Nathalie REDON	Maitre-Assistant, EMD, Douai

Spécialités doctorales :
SCIENCES ET GENIE DES MATERIAUX
MECANIQUE ET INGENIERIE
GENIE DES PROCEDES
SCIENCES DE LA TERRE
SCIENCES ET GENIE DE L'ENVIRONNEMENT
MATHEMATIQUES APPLIQUEES
INFORMATIQUE
IMAGE, VISION, SIGNAL
GENIE INDUSTRIEL
MICROELECTRONIQUE

Responsables :
K. Wolski Directeur de recherche
S. Drapier, professeur
F. Gruy, Maître de recherche
B. Guy, Directeur de recherche
D. Graillot, Directeur de recherche
O. Roustant, Maître-assistant
O. Boissier, Professeur
JC. Pinoli, Professeur
A. Dolgui, Professeur

EMSE : Enseignants-chercheurs et chercheurs autorisés à diriger des thèses de doctorat (titulaires d'un doctorat d'État ou d'une HDR)

AVRIL	Stéphane	PR2	Mécanique et ingénierie	CIS
BATTON-HUBERT	Mireille	PR2	Sciences et génie de l'environnement	FAYOL
BENABEN	Patrick	PR1	Sciences et génie des matériaux	CMP
BERNACHE-ASSOLLANT	Didier	PR0	Génie des Procédés	CIS
BIGOT	Jean Pierre	MR(DR2)	Génie des Procédés	SPIN
BILAL	Essaid	DR	Sciences de la Terre	SPIN
BOISSIER	Olivier	PR1	Informatique	FAYOL
BORBELY	Andras	MR(DR2)		SMS
BOUCHER	Xavier	PR2	Génie Industriel	FAYOL
BRODHAG	Christian	DR	Sciences et génie de l'environnement	FAYOL
BURLAT	Patrick	PR2	Génie Industriel	FAYOL
COLLOT	Philippe	PR0	Microélectronique	CMP
COURNIL	Michel	PR0	Génie des Procédés	DIR
DARRIEULAT	Michel	IGM	Sciences et génie des matériaux	SMS
DAUZERE-PERES	Stéphane	PR1	Génie Industriel	CMP
DEBAYLE	Johan	CR	Image Vision Signal	CIS
DELAFOSSSE	David	PR1	Sciences et génie des matériaux	SMS
DESRAYAUD	Christophe	PR2	Mécanique et ingénierie	SMS
DOLGUI	Alexandre	PR0	Génie Industriel	FAYOL
DRAPIER	Sylvain	PR1	Mécanique et ingénierie	SMS
FEILLET	Dominique	PR2	Génie Industriel	CMP
FOREST	Bernard	PR1	Sciences et génie des matériaux	CIS
FORMISYN	Pascal	PR0	Sciences et génie de l'environnement	DIR
FRACZKIEWICZ	Anna	DR	Sciences et génie des matériaux	SMS
GARCIA	Daniel	MR(DR2)	Génie des Procédés	SPIN
GIRARDOT	Jean-jacques	MR(DR2)	Informatique	FAYOL
GOEURLOT	Dominique	DR	Sciences et génie des matériaux	SMS
GRAILLOT	Didier	DR	Sciences et génie de l'environnement	SPIN
GROSSEAU	Philippe	DR	Génie des Procédés	SPIN
GRUY	Frédéric	PR1	Génie des Procédés	SPIN
GUY	Bernard	DR	Sciences de la Terre	SPIN
GUYONNET	René	DR	Génie des Procédés	SPIN
HAN	Woo-Suck	CR		SMS
HERRI	Jean Michel	PR1	Génie des Procédés	SPIN
INAL	Karim	PR2	Microélectronique	CMP
KLOCKER	Helmut	DR	Sciences et génie des matériaux	SMS
LAFOREST	Valérie	MR(DR2)	Sciences et génie de l'environnement	FAYOL
LERICHE	Rodolphe	CR	Mécanique et ingénierie	FAYOL
LI	Jean Michel		Microélectronique	CMP
MALLIARAS	Georges	PR1	Microélectronique	CMP
MOLIMARD	Jérôme	PR2	Mécanique et ingénierie	CIS
MONTHEILLET	Franck	DR	Sciences et génie des matériaux	SMS
PERIER-CAMBY	Laurent	PR2	Génie des Procédés	DFG
PIJOLAT	Christophe	PR0	Génie des Procédés	SPIN
PIJOLAT	Michèle	PR1	Génie des Procédés	SPIN
PINOLI	Jean Charles	PR0	Image Vision Signal	CIS
ROUSTANT	Olivier	MA(MDC)		FAYOL
STOLARZ	Jacques	CR	Sciences et génie des matériaux	SMS
SZAFNICKI	Konrad	MR(DR2)	Sciences et génie de l'environnement	CMP
TRIA	Assia		Microélectronique	CMP
VALDIVIESO	François	MA(MDC)	Sciences et génie des matériaux	SMS
VIRICELLE	Jean Paul	MR(DR2)	Génie des Procédés	SPIN
WOLSKI	Krzysztof	DR	Sciences et génie des matériaux	SMS
XIE	Xiaolan	PR1	Informatique	CIS

ENISE : Enseignants-chercheurs et chercheurs autorisés à diriger des thèses de doctorat (titulaires d'un doctorat d'État ou d'une HDR)

FORTUNIER	Roland	PR	Sciences et Génie des matériaux	ENISE
BERGHEAU	Jean-Michel	PU	Mécanique et Ingénierie	ENISE
DUBUJET	Philippe	PU	Mécanique et Ingénierie	ENISE
LYONNET	Patrick	PU	Mécanique et Ingénierie	ENISE
SMUROV	Igor	PU	Mécanique et Ingénierie	ENISE
ZAHOUANI	Hassan	PU	Mécanique et Ingénierie	ENISE
BERTRAND	Philippe	MCF	Génie des procédés	ENISE
HAMDI	Hédi	MCF	Mécanique et Ingénierie	ENISE
KERMOUCHE	Guillaume	MCF	Mécanique et Ingénierie	ENISE
RECH	Joël	MCF	Mécanique et Ingénierie	ENISE
TOSCANO	Rosario	MCF	Mécanique et Ingénierie	ENISE
GUSSAROV Andrey	Andrey	Enseignant contractuel	Génie des procédés	ENISE

PR 0	Professeur classe exceptionnelle	Ing.	Ingénieur
PR 1	Professeur 1 ^{ère} classe	MCF	Maître de conférences
PR 2	Professeur 2 ^{ème} classe	MR (DR2)	Maître de recherche
PU	Professeur des Universités	CR	Chargé de recherche
MA (MDC)	Maître assistant	EC	Enseignant-chercheur
DR	Directeur de recherche	IGM	Ingénieur général des mines

SMS	Sciences des Matériaux et des Structures
SPIN	Sciences des Processus Industriels et Naturels
FAYOL	Institut Henri Fayol
CMP	Centre de Microélectronique de Provence
CIS	Centre Ingénierie et Santé

NNT : 2012 EMSE 0677

Zaher AL BARAKEH

Suivi de pollution atmosphérique par système multi-capteurs – méthode mixte de classification et de détermination d'un indice de pollution.

Spécialité: Génie de Procédé

Mots clefs : Pollution atmosphérique, capteurs de gaz semi-conducteurs, réseaux de
capteurs, logique floue.

Résumé :

Cette thèse a pour objectif le développement d'un système multi-capteurs de gaz permettant une évaluation en continu et en temps réel des différentes types de pollution atmosphérique en zone urbain, en classifiant notamment les pollutions de type urbaine, photochimique, ou encore liée au trafic. Le projet se base sur l'utilisation de différents capteurs de gaz de type semi-conducteur disponibles dans le commerce qui sont intégrés dans un dispositif autonome et portable, afin qu'il puisse fonctionner sur site.

Dans un premier temps, et en grande partie à Saint-Etienne, différents types de capteurs sont sélectionnés puis leurs performances sont testées sur un banc simulant les atmosphères polluées et développé pour l'occasion. Afin de pallier aux problèmes de non répétabilité et de dérive de la ligne de base et de la sensibilité, des procédures de prétraitement de standardisation sont mises au point.

Dans un deuxième temps, et en grande partie à Douai, différents sites de tests sont identifiés et leurs historiques de pollution sont étudiés. Plusieurs campagnes en stations de mesure d'une semaine, recouvrant les différentes saisons et les différents types de sites, sont alors menées. Il y est collecté conjointement les signaux des capteurs et des analyseurs de gaz réglementés. Des méthodes basées sur les réseaux de neurones sont alors appliquées afin d'obtenir conjointement, à partir des signaux des capteurs, une classification parmi 3 types de pollutions (urbaine, trafic et photochimique) ainsi qu'un indicateur global de qualité de l'air. Ces méthodes utilisent une approche basée sur la logique floue afin d'éviter les problèmes d'effet de bord.

Remerciements :

Cette thèse doit beaucoup aux nombreuses personnes qui m'ont encouragé, soutenu et conforté au long de toutes ces années. Qu'elles trouvent dans ce travail l'expression de mes plus sincères remerciements.

*Tout d'abord, j'adresse mes chaleureux remerciements aux Co-directeurs de cette thèse, M. **Christophe PIJOLAT** et Mme. **Nadine LOCOGE**, pour m'avoir fait confiance malgré les connaissances plutôt légères que j'avais en octobre 2009 sur le domaine des capteurs de gaz et la pollution atmosphérique, puis pour m'avoir guidé, encouragé, conseillé, fait beaucoup voyager pendant presque trois ans tout en me laissant une grande liberté et en me faisant l'honneur de me déléguer plusieurs responsabilités dont j'espère avoir été à la hauteur. Je sais infiniment gré à vous de m'avoir prodigué maints conseils.*

*Pour son attention de tout instant sur mes travaux, pour ses conseils avisés et son écoute qui ont été prépondérants pour la bonne réussite de cette thèse. Son énergie et sa confiance qui ont été des éléments moteurs pour moi. J'adresse mes remerciements à mon Co-encadrant, M. **Philippe BREUIL** qui j'ai pris un grand plaisir à travailler avec lui.*

*Mes remerciements vont également à mon Co-encadrant Mme. **Nathalie REDON**, pour la gentillesse et la patience qu'elle a manifestée à mon égard durant cette thèse, pour tous ses conseils et ses recommandations, spécialement aux moments critiques où je manquais de repères.*

Je ne sais pas comment exprimer ma gratitude à ces personnes autrement qu'en leur promettant d'agir comme eux avec des étudiants dans ma situation, si un jour l'occasion m'en est donnée.

*Je voudrais remercier les rapporteurs de cette thèse Messieurs les Profs. **Jacques NICOLAS** et **Marcel BOUVET** pour l'intérêt qu'ils ont porté à mon travail dont ils ont accepté d'être les rapporteurs de cette thèse. Ils ont également contribué par leurs nombreuses remarques et suggestions à améliorer la qualité de ce mémoire, et je leur en suis très reconnaissant.*

*J'associe à ces remerciements Monsieur **Pierre LANTERI** et Madame **Laure CHEVALLIER** pour avoir accepté d'examiner mon travail je les en remercie profondément.*

*Ces remerciements seraient incomplets si je n'en adressais pas à l'ensemble des membres de personnel du **centre SPIN** de l'**Ecole des Mines de Saint Etienne**, ainsi que les personnels du **Département du Chimie et Environnement** de l'**Ecole des Mines de Douai**, ceux sans qui cette thèse ne serait pas ce qu'elle est. Pour les discussions que j'ai eu la chance d'avoir avec eux, leurs suggestions ou contributions, aussi bien pour leur soutien logistique et moral ainsi que pour la très bonne ambiance que j'ai toujours trouvée.*

*Je souhaite également remercier **le personnel d'Azimut Monitoring**, pour leur accueil chaleureux et pour les efforts qu'ils ont réalisé afin que j'achève mon travail dans les meilleures conditions possibles.*

*Enfin, une pensée émue pour tous les étudiants avec qui j'ai partagé une salle, un café, un repas ou une console d'ordinateur pendant ces trois années : **Damien, Pauline, Mounir, Kevin, Florent, Aurélie, Pedro, Dimitry, Ivan, Mokhtar, Jing, Carole, Geoffroy, Jinan** ... Et finalement **Raquel**, merci de m'avoir épaulé et encouragé pendant ces longs mois de rédaction, aussi que dans mes moments de doutes et de craquages. Ton soutien a été précieux et je n'ai jamais eu l'occasion de te le dire et de t'en remercier. Je te dois beaucoup...*

Mes remerciements s'adressent enfin à mon papa, ma maman, et mon frère, pour leur soutien et leur patience sans lesquels je n'en serais pas là aujourd'hui. Je sais que mon absence a été longue et j'espère pouvoir un jour rattraper le retard accumulé. Vous avez toujours ma gratitude éternelle.

Introduction	1
Chapitre I :	
La pollution atmosphérique : les différents polluants, les méthodes réglementaires de surveillance, les méthodes alternatives.....	7
1. La pollution atmosphérique	8
1.1. Les polluants atmosphériques et leurs sources	9
1.1.1 Les polluants primaires	10
1.1.2 Les polluants secondaires.....	17
1.2. Réglementation.....	19
2. Mesures et méthodes normalisées.....	23
2.1. Différentes méthodes de mesures spécifiques aux polluants.....	23
2.2. Réseaux de surveillance de la qualité d'air.....	24
2.3. Stations de mesure de la pollution.....	26
2.4. Information du public – Indice ATMO	29
2.5. Conclusion.....	30
3. Les capteurs de gaz	31
3.1. Les différents types de capteurs de gaz	32
3.1.1. Les capteurs à électrolyte solide.....	33
3.1.2. Capteurs capacitifs	35
3.1.3. Capteurs résistifs, capteurs à oxydes semi-conducteurs (SnO ₂)	36
3.1.4. Autres types de capteurs - capteurs physiques	41
3.2. Utilisation des capteurs dans la qualité de l'air	42
3.3. Caractéristiques métrologiques d'un capteur de gaz à s-c.....	42
3.3.1. Sensibilité.....	44
3.3.2. Sélectivité.....	44
3.3.3. Stabilité.....	45
3.4. Réseau des capteurs de gaz - Nez électroniques.....	46
3.4.1. Approches de mesures quantitatives ou qualitatives.....	47
3.4.2. Architecture d'un système multi-capteur	48
4. Analyse multi-variable.....	50
4.1. Principes de modélisation multi-variables.....	50
4.2. Problème de sur-modélisation :	52
4.3. Les différentes méthodes de modélisation.....	53
4.3.1. Méthodes linéaires simples	53

4.3.2.	Méthodes linéaires complexes, dites « factorielles »	54
4.3.3.	Les réseaux de neurones artificiels.....	57
4.4.	Système multi-capteurs et analyse multi-variables	62
4.4.1.	Système multi-capteurs et réseaux de neurones	62
4.4.2.	Prétraitement et normalisation des données	62
4.4.3.	Systèmes multi-capteurs et qualité de l'air.....	64
5.	Conclusion et objectifs de la thèse	67
Chapitre II :		
Développement, mise en œuvre et validation d'un réseau de capteurs.		71
1.	Introduction.....	71
2.	Choix des capteurs	72
2.1.	Propriétés des capteurs	73
2.2.	Mise en œuvre des capteurs à semi-conducteurs.....	75
2.3.	Systèmes de mesure d'humidité et de température du gaz	76
2.4.	Conception de la cellule de capteurs en laboratoire	80
3.	Développement de bancs de test	80
3.1.	Banc de test de validation	81
3.1.1.	Génération des concentrations gazeuses de polluants.....	82
3.1.2.	Système d'humidification du mélange gazeux à l'aide du banc de test.....	84
3.2.	Banc de test développé pour les mesures « terrain »	86
4.	Réponses des capteurs.....	87
4.1.	Mesure de la conductance des capteurs	87
4.2.	Effet de la température et de l'humidité sur les capteurs.....	89
4.2.1.	Effet de la température sur les capteurs	91
4.2.2.	Influence de l'humidité absolue sur la réponse des capteurs	92
4.3.	Réponse des capteurs de gaz – signature des gaz	95
4.3.1.	Cas des gaz réducteurs	95
4.3.2.	Cas des gaz oxydants.....	97
4.3.3.	Signature des différents gaz sur les différents capteurs	98
5.	Gestion des écarts de sensibilité et de la dérive de la ligne de base des capteurs.....	100
5.1.	Evolution à long terme des lignes de base.....	101
5.2.	Test de calibrage de G_0	102
5.3.	Implémentation d'une standardisation de sensibilité.....	105
5.4.	Effet de la standardisation sur les sensibilités mesurées	107

5.5. Dérive temporelle des sensibilités	109
6. Conclusion du chapitre	110
Chapitre III :	
Classification et indexation de la pollution de l'air avec un réseau de capteurs.	113
1. Tests terrain.....	114
2. Caractéristiques des pollutions atmosphériques des sites étudiés.....	116
2.1. Analyse des concentrations des polluants par type de station	116
2.2. Identification de signatures.....	120
2.2.1. Les formes journalières	120
2.2.2. Les formes hebdomadaires :.....	124
3. Démarches de modélisation-prédiction.....	126
3.1. Prétraitements de données	126
3.2. Elimination des points aberrants.....	128
3.3. Choix des variables.....	130
3.4. Apprentissage des modèles, limites de validité	131
4. Comportement des capteurs	134
4.1. Etude des corrélations signaux des capteurs - polluants.....	134
4.2. Comportement de type oxydant ou réducteur des capteurs	137
4.3. Tentative de prédiction quantitative	138
4.4. Prédiction catégorielle	140
4.4.1. Cas de la pollution trafic	140
4.4.2. Cas de la pollution urbaine	141
4.4.3. Cas de la pollution photochimique :.....	142
4.4.4. Démarche de catégorisation :	144
4.5. Catégorisation avec la logique floue	148
4.5.1. Choix d'une fonction non linéaire.....	149
4.5.2. Application de la logique floue à la catégorisation des pollutions.....	151
4.5.3. Résultat de catégorisation.....	153
4.6. Prédiction mixte catégorielle et indicielle	156
5. Résultats de prédiction mixte.....	159
5.1. Optimisation des modèles.....	159
5.1.1. Normalisation et standardisation des sensibilités	160
5.1.2. Choix des variables	161
5.1.3. Seuils de catégorisation	161

5.2.	Modèles mono-cellule et mono-site, modèles locales.	162
5.3.	Portabilité des modèles	165
5.3.1.	Modèles mono-cellule et inter-site	165
5.3.2.	Modèles mono-cellule et multi-site.....	166
5.3.3.	Modèles inter-cellules.	167
5.4.	Amélioration du critère de performance.....	170
5.5.	Sensibilité des modèles aux variables	171
5.6.	Validité des prédictions	172
6.	Conclusion du chapitre	175
	Conclusion générale	177
	Bibliographie.....	181
	Publications et communications.....	193
	Annexes :	195
	Annexe A : Systèmes multi-capteurs développées :	195
	Annexe B : Corrélations entre les concentrations des polluants :	198
	Annexe C : Forme journalière et hebdomadaire de pollution :	205
	Annexe D : Etude des comportements des capteurs :	211

Introduction

Dès la fin du XVIII^{ème} siècle, en raison des problèmes causés par la combustion du charbon dans les grandes villes (Hadfield, et al., 1999), il est apparu nécessaire de mettre en place un système de management de la qualité de l'air. A l'époque, à cause des limites technologiques et en l'absence d'instrumentation adaptée à la mesure de cette pollution, il n'existait aucune réglementation. Depuis lors, entre industrialisation exponentielle et augmentation drastique de l'activité anthropique, la problématique de la pollution de l'air n'a cessé de croître. Compte-tenu des progrès technologiques réalisés en matière de mesures d'espèces gazeuses dans l'atmosphère, avec des limites de détections (LD) atteignant quelques particules par trillion (10^{12}) (ppt) (Koga, et al., 2001) des dizaines de normes et directives ont vu le jour et réglementent désormais nombre de polluants dans l'atmosphère (Ramanathan, et al., 2009). Certaines de ces espèces peuvent affecter notre environnement naturel, comme les pluies acides, liées aux émissions d'oxyde de soufre, qui détruisent la végétation, ou avoir un impact négatif sur la santé humaine allant de la simple irritation (Llop, et al., 2008) à des allergies comme l'asthme (Peden, 1996) ou encore déclencher des maladies graves aux conséquences létales (Pascal, 2009).

L'étude de la pollution atmosphérique vise à bien en comprendre ses particularités dans le but de limiter au mieux ses effets (Dimo, et al., 2010). De nombreux efforts ont été faits pour contrôler et réduire les émissions des polluants provenant du transport, de l'industrie et de l'agriculture et cela a eu un impact remarquable sur les taux de dioxyde de soufre et monoxyde de carbone dans l'atmosphère : ceux-ci, après avoir présenté des quantités croissantes durant la première moitié du XX^{ème} siècle, ont vu leurs teneurs réduites de près d'un tiers sur les deux dernières décennies. A titre d'exemple, les émissions de dioxyde de soufre en France sont passées de 1326 Gg en 1990 à 303 Gg en 2009¹. Sur la même période, les émissions de monoxyde de carbone, quant à elles, ont diminué de 10890 Gg à 3950 Gg. En revanche, la situation se dégrade pour d'autres polluants, comme les oxydes d'azote et l'ozone dont les niveaux dépassent souvent les seuils d'alerte environnementaux dans les grandes villes. L'ozone, qui avait des niveaux moyens de fond compris entre 10 et 15 ppb au début du siècle, atteint maintenant 30 à 40 ppb, en de nombreux endroits dans le monde (Percy, et al., 2004).

Afin de répondre à ces problématiques environnementales, des organismes nationaux et internationaux dédiés à la mesure en continu des différents types de pollution ont été créés. Leur but étant de vérifier les niveaux par rapport aux standards de qualité environnementale exigés. Le suivi de la pollution atmosphérique se fait donc actuellement via des réseaux de mesure, dont les données s'appuient sur des analyseurs industriels placés dans des stations fixes réparties judicieusement sur le territoire à surveiller. Avec ces stations de mesures, il est possible d'effectuer un suivi continu et précis des teneurs des polluants réglementés sur des lieux stratégiques. En revanche, les mesures effectuées par ces stations ne répondent pas

¹ *Données obtenues du Centre on Emission Inventories and Projection*
<http://www.ceip.at> <http://www.ceip.at/emission-data-webdab/officially-reported-emission-data/>

toujours aux applications qui s'intéressent à l'étude qualitative de l'atmosphère et nécessitent une mobilité géographique (Tsujita, et al., 2005).

Une méthode alternative au suivi de la pollution atmosphérique réside donc en l'utilisation de systèmes multi-capteurs de gaz (Thery-Merland, et al., 2006) basés sur des capteurs de gaz chimiques et/ou à semi-conducteurs (Ma, et al., 2008). Ces réseaux de capteurs visent à compléter le panel de l'instrumentation actuellement disponible en offrant la possibilité de tracer rapidement une signature de pollution sans nécessairement faire une étude exhaustive par analyse physico-chimique standard de la composition de l'atmosphère qui la constitue.

Le suivi spatio-temporel de la pollution atmosphérique est un domaine où l'utilisation de ce type de systèmes multi-capteurs constitue un outil complémentaire par rapport aux techniques lourdes d'analyse, de par leurs faibles coûts et leur encombrement réduit. Ils permettent notamment de développer des réseaux de mesure plus simples, faciles à mettre en œuvre et mobiles, pour effectuer un maillage géographique d'une zone impactée par une pollution spécifique. Habituellement, ce type de réseaux de capteurs est dédié à la qualification de nuisances olfactives comme en témoignent les développements récents de « nez électroniques ». Nous proposons pour notre part d'utiliser un système multi-capteurs pour la mesure de la pollution atmosphérique au sens large, sujet rarement abordé par ce type de dispositifs. L'objectif fixé dans ce travail de thèse est donc la conception d'un système portable autonome, basé sur des capteurs de gaz commerciaux, pour la détection et la classification de différents types et niveaux de pollutions atmosphériques pouvant survenir sur un fond de pollution urbaine ou photochimique. L'étude proposée se démarque résolument des précédentes travaux menés au laboratoire de l'EMSE sur des systèmes multi-capteurs (en particulier les thèses de N. Perdreau (Perdreau, 2000) et M. Kamionka (Kamionka, 2005) par notre volonté d'utiliser des capteurs disponibles commercialement afin de maîtriser au mieux

les problèmes de reproductibilité et de dérive des capteurs, ainsi que de préfigurer d'éventuels transferts de technologies.

Dans un premier temps, nous proposons de familiariser le lecteur avec la problématique de la pollution atmosphérique ainsi que les différentes méthodes et réglementations associées à sa mesure. Nous poursuivons cette étude bibliographique par une description des avantages et inconvénients liés à l'utilisation de capteurs de gaz, pour la mesure de la pollution de l'air. Les différents éléments physiques, ainsi que les modèles mathématiques associés constituant les systèmes à base de réseaux de capteurs, sont ensuite présentés, accompagnés d'un panorama des problèmes habituellement rencontrés lors de qualifications réelles sur le terrain de la pollution atmosphérique.

Dans un second temps, un chapitre est consacré à la méthodologie choisie pour concevoir trois systèmes identiques de détection des gaz à base de réseaux de capteurs. Les bancs développés pour la caractérisation des performances métrologiques de ces systèmes y sont exposés, ainsi que la mise en œuvre de nouvelles méthodes de traitement de leur signal permettant de s'affranchir de certains défauts récurrents, comme les problèmes de dérive, ou de reproductibilité.

Enfin, les principaux travaux de mesure et d'analyse de pollution réalisés sur sites réels en association avec les associations de surveillance de la qualité de l'air sont détaillés dans le dernier chapitre. Des outils de traitement des résultats de mesure faisant appel aux réseaux de neurones combinés à de la logique floue ont été pour la première fois appliqués à des données issues de réseaux de capteurs pour de la qualification de pollution atmosphérique.

Ce travail est le fruit d'une collaboration entre l'EMSE² et l'ENSM-Douai³. Les expériences acquises par le département MICC au centre SPIN de l'EMSE dans le domaine

² Ecole Nationale supérieure des Mines de Saint Etienne

des capteurs chimiques de gaz, spécialement en ce qui concerne les capteurs à semi-conducteurs, couplées avec les expériences du département Chimie et Environnement de l'Ecole des Mines de Douai dans le domaine de la mesure et de la surveillance de la qualité de l'air, donnent des bases solides à ce sujet. Ce travail fait partie du programme « GEM-Air » qui regroupe les compétences de 4 écoles des Mines dans le domaine de l'étude de la qualité d'air, et ceci dans le cadre plus général des projets collaboratifs du GEM, Groupe des Ecoles des Mines.

³ Ecole Nationale Supérieure des Mines de Douai

Chapitre I :

La pollution atmosphérique : les différents polluants, les méthodes réglementaires de surveillance, les méthodes alternatives.

L'objectif de ce chapitre est de présenter globalement la problématique de la pollution atmosphérique et du suivi des concentrations des polluants gazeux afin de dresser un bilan des travaux pouvant être susceptibles d'apporter des éléments exploitables par l'utilisation de réseaux de capteurs. Dans cette optique, nous décomposons ce chapitre de la façon suivante :

La première partie est dédiée à la description de la pollution atmosphérique, les polluants gazeux ainsi que leurs caractéristiques. Ensuite, nous présentons les différentes méthodes pouvant être utilisées pour la mesure normalisée de ces polluants dans le cadre de leur suivi réglementaire.

La deuxième partie est consacrée aux différents capteurs de gaz et à leur utilisation comme méthode non réglementaire pour la mesure des polluants gazeux.

Finalement nous décrivons quelques méthodes d'analyse multi-variables habituellement utilisées en association avec des systèmes multi-capteurs de gaz pour la détection des espèces gazeuses. Le contexte ainsi bien établi, permet de montrer les aspects originaux de notre travail et par conséquent de bien poser les objectifs visés dans le cadre de cette thèse.

1. La pollution atmosphérique

La pollution atmosphérique peut être définie par la présence de polluants (gazeux ou particules) dans l'atmosphère, pouvant provoquer des effets nocifs sur l'environnement et la santé (Arquès, 1998). Les sources de cette pollution peuvent être soit naturelles (feu de forêt, éruption volcanique,...), soit anthropiques, c'est-à-dire liées à l'activité humaine (Mayer, 1999). Dans ce dernier cas, la pollution est souvent le résultat direct des progrès industriels de ces derniers siècles, comme par exemple l'émission continue et parfois sans précaution des polluants associés aux processus de combustion (véhicules automobiles, installations industrielles, production d'énergie par combustion...). Depuis quelques dizaines d'années, des études ont montré un lien entre la dégradation de l'environnement et de la santé humaine, et la présence de ces polluants dans l'atmosphère (Monk, et al., 2009) (Anderson, 2009) (Kulkarni, et al., 2008) (FNORS, Septembre 2008). Ainsi, la pollution atmosphérique est considérée comme étant responsable d'environ 800 000 décès prématurés chaque année dans le monde (OECD, 2008). En exemple, la figure 1-1 présente l'évolution prévisionnelle des décès prématurés liés à l'excès d'ozone entre 2000 et 2030 (OECD, 2008). Quelle que soit la zone géographique, la situation est déjà très inquiétante avec près de 8 à 15 décès prématurés imputés à la pollution atmosphérique par million d'habitants et par an, mais les prévisions à 20 ans sont très pessimistes avec en particulier une augmentation d'un facteur 10 dans la zone Asie.

L'accumulation de ces polluants est aussi à l'origine des phénomènes de pluies acides qui ont un impact très négatif sur la végétation ou encore du réchauffement terrestre (Arquès, 1998). Afin de bien cerner le contexte dans lequel ce travail se situe et de bien définir le cahier des charges de l'outil de diagnostic que nous proposons de développer, nous avons donc recensé dans le paragraphe suivant les polluants atmosphériques primaires ou secondaires, leurs sources, et la réglementation associée à l'échelle européenne le cas échéant.

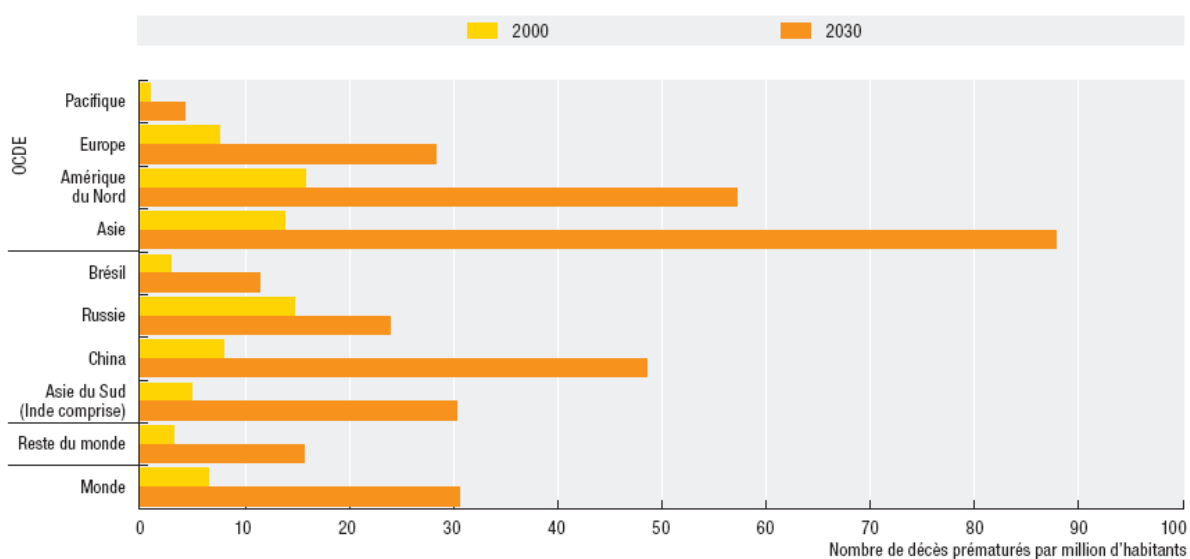


Figure 1-1 : décès prématurés imputables à l'exposition à l'ozone dans les zones urbaines en 2000 et 2030 (OECD, 2008).

1.1. Les polluants atmosphériques et leurs sources

Les sources de polluants atmosphériques peuvent être divisées en 2 catégories : les sources anthropiques et les sources biogéniques (Arquès, 1998). Par définition, une source anthropique est une source de pollution liée à l'activité humaine ; elle se subdivise en 3 catégories : secteur industriel, secteur résidentiel et transport. Les sources biogéniques sont les sources associées aux émissions d'origines naturelles liées à la biosphère (ORS, décembre 2007). L'agriculture est également une source de pollution majeure avec en particulier les engrais azotés ou encore l'élevage des animaux à l'origine de pollution par le méthane (produit par leur système digestif), et par l'ammoniac (issu de la transformation de l'urée).

Enfin, on intègre également dans les sources biogéniques les phénomènes naturels comme les nuages de cendres associées aux éruptions volcaniques (Arquès, 1998).

Le caractère dangereux de ces polluants, a conduit l'Union Européenne à légiférer : des directives et des normes ont été mises en place pour limiter et contrôler les émissions polluantes des principales sources anthropiques et également permettre la surveillance des polluants réglementés avec des objectifs de qualité uniforme à l'échelle européenne.

1.1.1 Les polluants primaires

Les polluants peuvent être classés comme primaires ou secondaires. Par définition, les polluants primaires sont ceux qui proviennent directement d'une source polluante identifiée, tels que le monoxyde de carbone provenant des processus de combustion et présent dans les gaz d'échappement des voitures, le dioxyde de soufre provenant de l'industrie (Mayer, 1999)... Les polluants secondaires ne sont pas émis directement par des sources connues et identifiées. Au contraire, ils se forment dans l'air à la suite de réactions mettant en jeu les composés émis par les sources primaires, notamment lors d'épisodes météorologiques particuliers. Un exemple majeur de polluant secondaire est l'ozone troposphérique, un des nombreux polluants secondaires qui forment le smog photochimique, mais on peut citer aussi les NPA (nitrates de peroxyacétyl) ou encore les COV (composé organiques volatiles) dits « COV secondaires ».

Le monoxyde de carbone (CO) : il résulte de la combustion incomplète des combustibles fossiles (charbons, fuels, ...). Ce gaz est toxique (pour les êtres vivants) et peut causer des asphyxies mortelles à hautes concentrations. Il est considéré comme l'un des composés principaux dans la problématique de la qualité de l'air intérieur, l'AFSSET⁴ a introduit une

⁴ AFSSET : Agence Française de Sécurité Sanitaire de l'Environnement et du Travail

valeur guide de 100 mg/m^3 pour une exposition de 15 min en CO comme une valeur limite de cette espèce au sens de la qualité de l'air intérieur. Pour ce qui concerne l'air extérieur, la Directive Européenne 2008/50/CE indique une valeur limite de 10 mg/m^3 pour un maximum journalier sur 8 heures (C.E., 2008). On trouve ce gaz en abondance dans les villes et aux abords des autoroutes à cause des véhicules à moteurs à combustion. Cela en fait un des traceurs de la pollution atmosphérique de type trafic.

Le dioxyde de soufre (SO_2) : ce gaz est produit naturellement par les volcans, il peut également être produit par différents procédés industriels comme les industries métallurgiques et les raffineries de pétrole. Ce gaz, toxique pour tous les êtres vivants, aide à la création des particules (aérosols sulfuriques) qui empêchent les rayons solaires d'atteindre la terre, ainsi il joue un rôle de refroidissant de la planète (Pascal, 2009). Conformément aux directives européennes, une surveillance des principales installations industrielles a été mise en place au cours des années 80 et la mise en œuvre de disposition de réduction à la source a permis une baisse drastique des émissions du SO_2 dans l'atmosphère. La figure 1-2 présente les différentes sources d'émissions de SO_2 enregistrées pendant l'année 2010, nous remarquons que 84% des émissions sont produites par la production d'énergie et le secteur industriel.

Ainsi, compte tenu de l'ensemble des réglementations mises en place, des concentrations qui restent relativement faibles de ce gaz peuvent être détectées aux alentours de certains sites industriels. La figure 1-3 montre les évolutions des moyennes annuelles de SO_2 en France dans les différents sites urbains et industriels. Nous remarquons d'une part que les concentrations ambiantes sont en baisse constante depuis les années 2000 et d'autre part que les concentrations moyennes en SO_2 sont toujours plus élevées en sites industriels qu'en sites urbains. Ce gaz peut ainsi être considéré comme traceur potentiel des émissions d'origine industrielle.

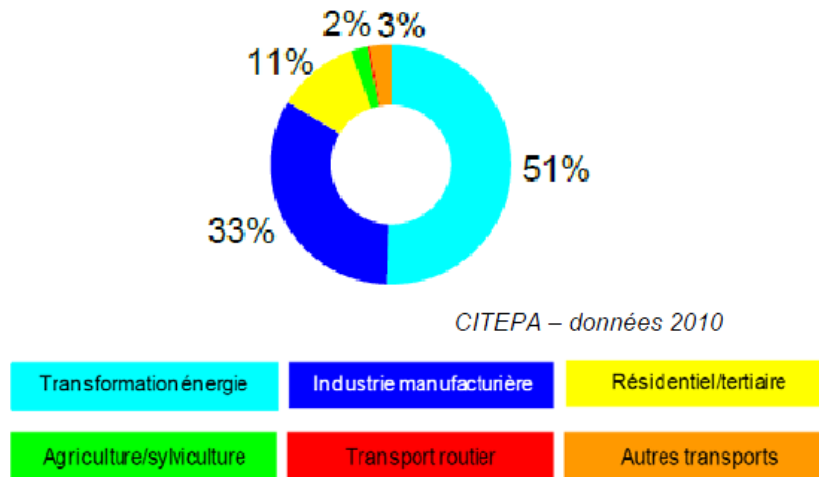
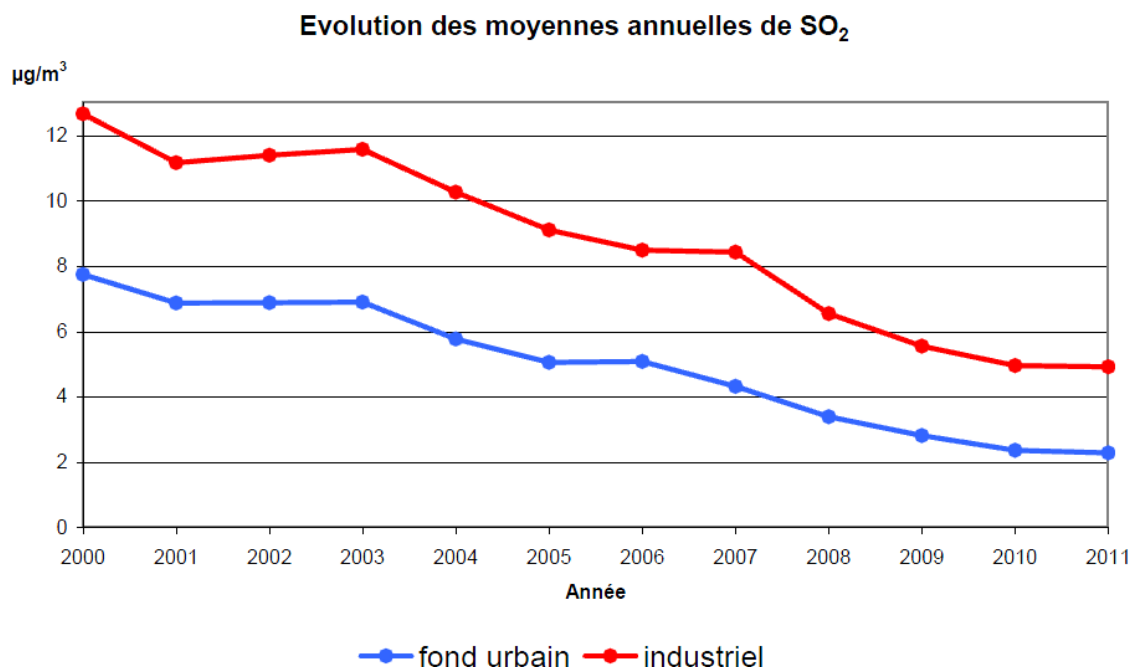


Figure 1-2 : Sources d'émission de SO₂ durant l'année 2010. Source : CITEPA⁵



Source : Base de données nationale de la qualité de l'air

Figure 1-3 : Evolution des moyennes annuelles de SO₂ en France dans des sites urbaines et industrielles. Source: BDQA⁶

Les oxydes d'azote (NO_x) : ils comprennent notamment le monoxyde d'azote (NO) et le dioxyde d'azote (NO₂). Ce sont des gaz fortement toxiques qui résultent de l'oxydation de l'azote de l'air par l'oxygène à température élevée, phénomène qui se produit généralement

⁵ Centre Interprofessionnel Technique d'Etude de la Pollution Atmosphérique
<http://www.citepa.org/fr/pollution-et-climat/polluants/aep-item/dioxyde-de-soufre>

⁶ Base de données nationale de la qualité de l'air :
<http://www.developpement-durable.gouv.fr/BDQA-Base-de-donnees-de-la-qualite.html>

lors des processus de combustion en particulier dans les moteurs à combustion interne et les centrales thermiques (Arquès, 1998). C'est dans les zones ayant une forte circulation automobile que nous observons de fortes concentrations en NO et NO₂. De plus, NO peut réagir avec l'oxygène de l'air pour se transformer en NO₂. La figure 1-4 montre les principales sources d'émission des NO_x enregistrées pendant l'année 2010, nous remarquons que 55% des émissions des NO_x ont pour origine le transport routier. Les NO_x pourraient donc constituer des traceurs de la pollution urbaine à l'échelle globale et plus ponctuellement de foyers importants de combustion comme les centrales de combustion urbaines et les trafics routiers.

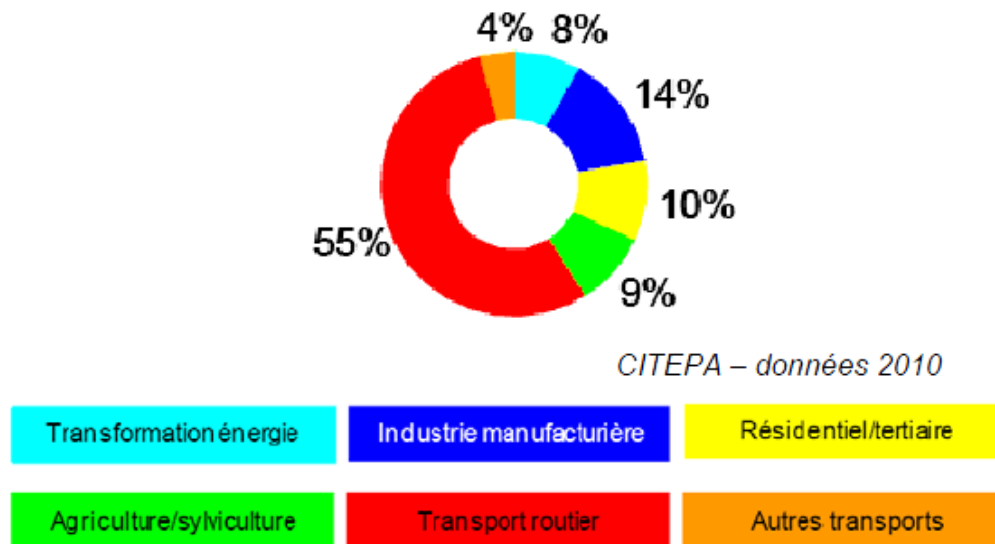


Figure 1-4: Sources d'émission des NO_x enregistrées en France durant l'année 2010. Source : CITEPA⁷

Les composés organiques volatiles (COV) : leur définition est spécifiée dans la Directive Européenne 2008/50/CE. Il s'agit des « composés organiques provenant de sources anthropiques et biogènes, autres que le méthane, capables de produire des oxydants photochimiques par réaction avec des oxydes d'azote sous l'effet du rayonnement solaire ». Une autre définition a été introduite par la directive 1999/13/CE: « tout composé organique, contenant au moins l'élément carbone et un ou plusieurs des éléments suivants : hydrogène,

⁷ <http://www.citepa.org/fr/pollution-et-climat/polluants/aep-item/oxydes-d-azote>

halogène, oxygène, soufre, phosphore, silicium ou azote à l'exception des oxydes de carbone et des carbonates et bicarbonates inorganiques, ayant une pression de vapeur de 0,01 kPa ou plus à une température de 293,15 K ou ayant une volatilité correspondante dans les conditions d'utilisation particulières ». L'émission anthropique des COV dans l'atmosphère provient de différentes sources, comme la combustion incomplète dans des moteurs ou des centrales thermiques, l'évaporation de produits raffinés, l'usage de solvants dans l'industrie, les peintures... La figure 1-5 présente les principales sources anthropiques des émissions des COV non méthaniques dans l'atmosphère, enregistrée entre les années 1988 et 2011.

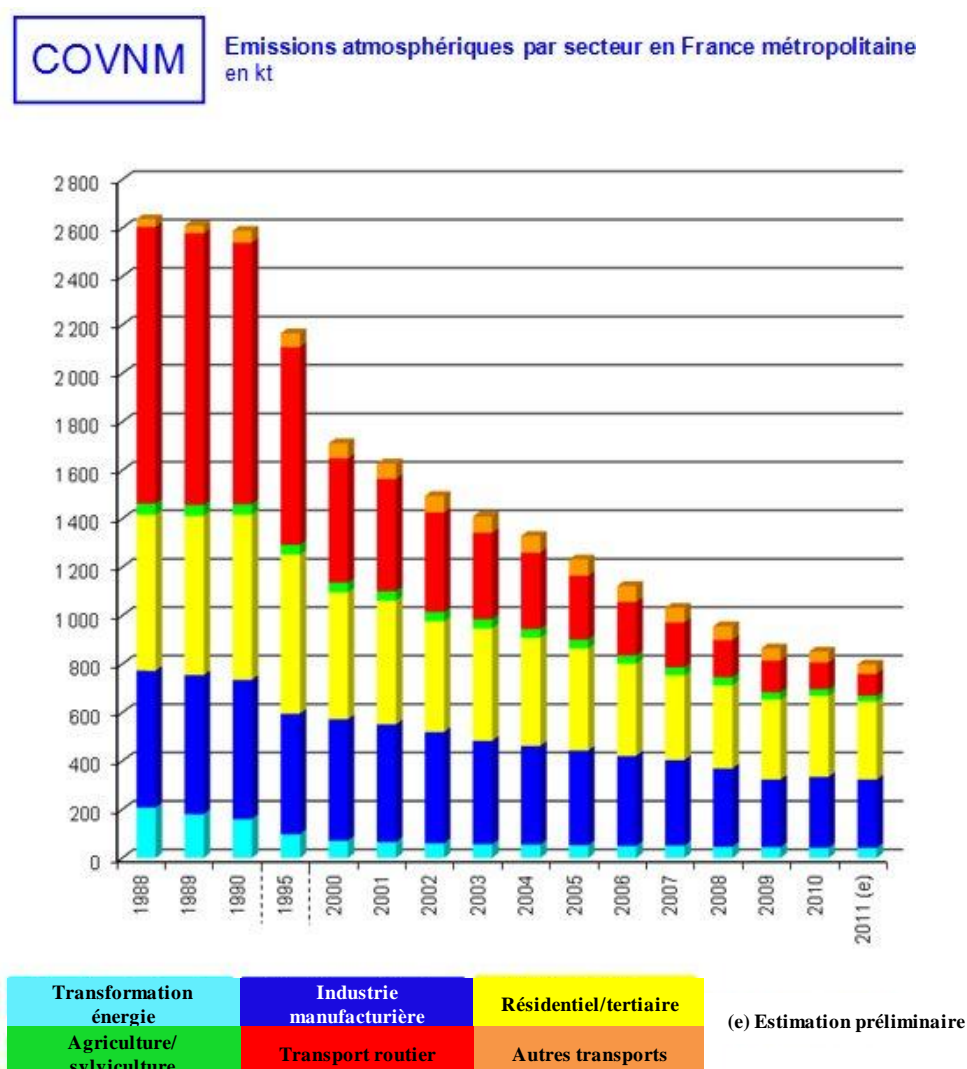


Figure 1-5 : Emission par secteur entre les années 1988 et 2011 des COVNM en France. Source: CITEPA⁸.

⁸ <http://www.citepa.org/fr/pollution-et-climat/polluants/aep-item/composes-organiques-volatils>

On remarque la forte baisse des émissions de ces polluants provenant des sources associées au transport routier essentiellement liées aux contrôles et réglementations plus exigeantes en termes d'émissions dans le secteur du transport. Actuellement les sources les plus importantes des émissions des COV sont les secteurs résidentiels et l'industrie. Il existe également des sources biogéniques à l'origine des émissions de COV (Arquès, 1998). Les COV biogéniques représentent plus que 90% des COV à l'échelle globale, mais en zone urbaine, cette part peut devenir négligeable. Ces COV participent à des réactions photochimiques dans la basse atmosphère, causant ainsi l'augmentation de la concentration en ozone dans la troposphère, ayant un effet indirect sur la santé (Le Cloirec, 2004). Ces COV peuvent être des traceurs des différents types de pollutions tant urbaines qu'industrielles (Liu, et al., 2012).

Les particules en suspension : ces particules (notées « PM » en anglais pour « *Particulate Matter* ») sont d'une manière générale les fines particules solides ou liquides en suspension dans l'air. Ces particules sont définies dans la directive 1999/30/CE comme « les particules passant dans un orifice d'entrée calibré avec un rendement de séparation de 50% pour un diamètre aérodynamique de 10 μm (cas des PM_{10}) ou de 2,5 μm (cas des $\text{PM}_{2,5}$) ». Ces particules proviennent de sources naturelles comme les éruptions volcaniques, la végétation (pollens...), les incendies de forêts... ou de sources anthropiques comme les émissions industrielles, la combustion des fossiles combustibles... On distingue 4 types de particules (C.E., 2008):

- **PM_{10}** particules en suspension dans l'air, d'un diamètre aérodynamique (ou diamètre aéraulique) inférieur à 10 μm ;
- **$\text{PM}_{2,5}$** dont le diamètre est inférieur à 2,5 μm , appelées « particules fines » ;
- **PM_1** dont le diamètre est inférieur à 1,0 μm , appelées « particules très fines » ;
- **$\text{PM}_{0,1}$** dont le diamètre est inférieur à 0,1 μm , appelées « particules ultrafines » ou nanoparticules.

Toutes ces particules sont dangereuses tant pour la santé des êtres humains, que pour la végétation. Par contre, seules les particules PM_{10} et $PM_{2,5}$ sont réglementées et ciblées par les Directives Européennes. L'augmentation des concentrations en particules dans l'air est corrélée avec l'augmentation des maladies cardiovasculaires, des allergies, des cancers du poumon, ce qui conduit à une diminution de l'espérance de vie (Pascal, 2009). La figure 1-6 montre les différentes sources d'émission des PM enregistrées en 2010. Les sources industrielles et résidentielles sont les producteurs majeurs des émissions des PM dans l'atmosphère.

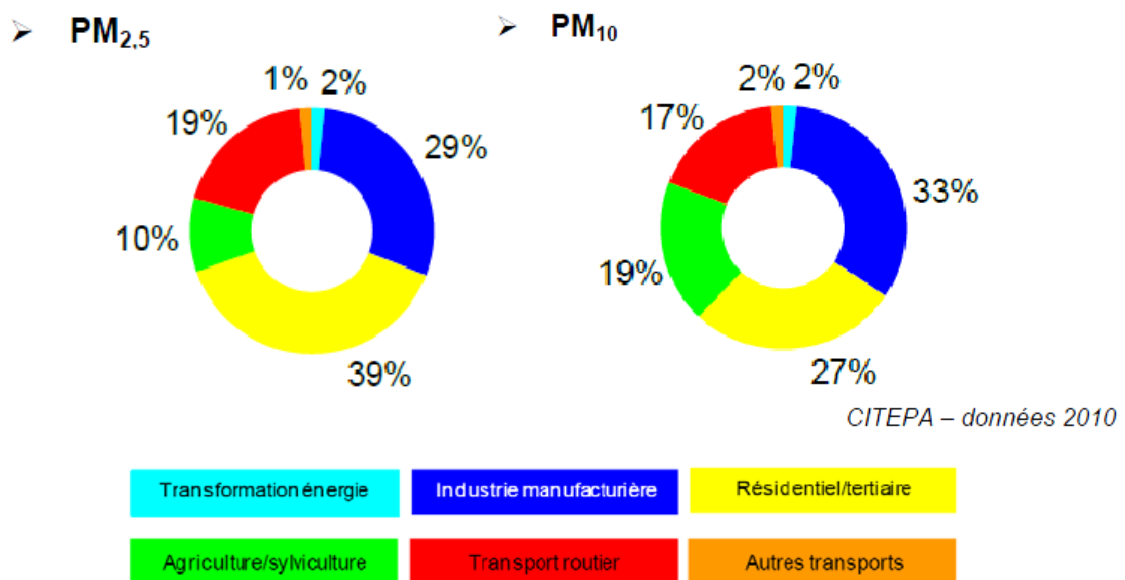


Figure 1-6 : Sources principales des émissions des PM en France durant l'année 2010. Source : CITEPA⁹

Le Méthane : ce gaz n'est pas réglementé en tant que polluant dans l'atmosphère par les Directives Européennes, mais peut être un bon traceur de certaines pollutions anthropiques. On le trouve à l'état naturel produit par des organismes vivants. Il constitue l'essentiel du gaz naturel qui est exploité comme combustible fossile. Il est l'un des principaux gaz à effet de serre et a un rôle important dans le réchauffement climatique. Depuis l'ère industrielle, la concentration de méthane n'a cessé d'augmenter dans l'atmosphère. Elle est passée de 700

⁹ <http://www.citepa.org/fr/pollution-et-climat/polluants/poussieres-en-suspension>

ppbv au début du 19^{ème} siècle à 1750 ppbv aujourd'hui en raison de l'augmentation des activités humaines liées à l'industrialisation et à l'urbanisation (figure 1-7).

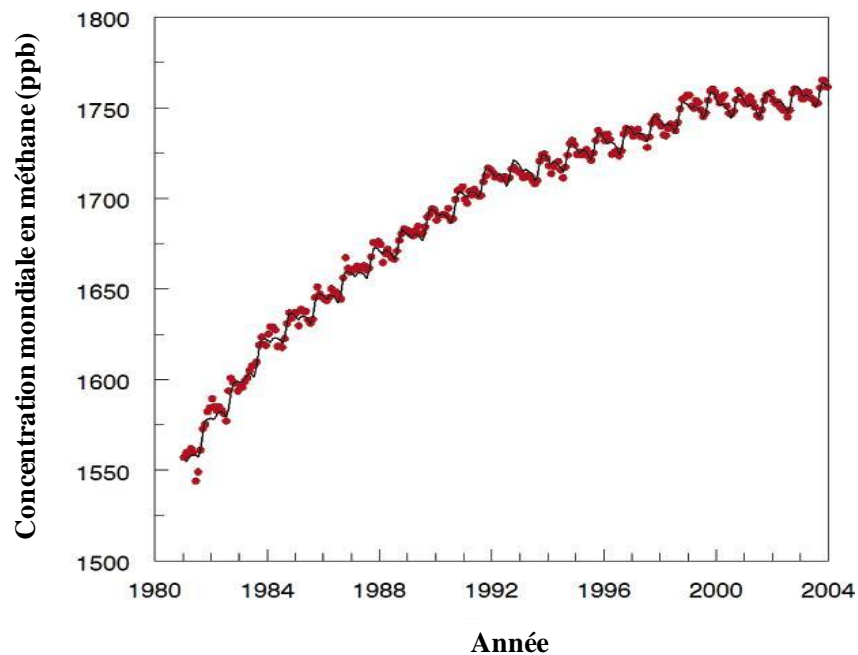


Figure 1-7 : Augmentation de la concentration du méthane dans le monde durant les 30 dernières années.
Source :OMER7¹⁰

1.1.2 Les polluants secondaires

Ce sont les polluants présents dans l'atmosphère qui se forment à la suite de réactions mettant en jeu les composés émis par des sources primaires ou encore qui ne proviennent pas d'une source directement polluante. Leur présence dans l'atmosphère est difficile à réguler car leur existence ne provient pas directement des activités anthropiques.

L'Ozone : dans la basse atmosphère et sous des conditions climatiques particulières (comme une augmentation de l'irradiation solaire), en zone urbaine, les émissions de précurseurs sont importantes : les hydrocarbures et les oxydes d'azote sont donc présents en grande quantité. Beaucoup d'ozone se formera mais, celui-ci réagira très vite, une fois formé, avec le monoxyde d'azote dont l'apport est constant. L'ozone sera donc en partie consommé au niveau des agglomérations et généralement présent en faible quantité.

¹⁰ Outils nuMERiques pour l'étude des Sciences de l'Environnement Terrestre : omer7.sedoo.fr

Comme pour tous les polluants secondaires, il est difficile de limiter la production d’ozone à partir de la mise en œuvre de limitations des émissions. Bien que la concentration globale des autres polluants dans l’atmosphère tende à baisser, celle de l’ozone reste intacte ou parfois subit de légères augmentations au cours des dernières années comme nous le montre la figure 1-8. C’est dans les zones rurales proches des agglomérations que les concentrations en O_3 enregistrées sont généralement les plus élevées. Ces concentrations plus élevées s’expliquent par la cinétique du formation de ce polluant secondaire lors du vieillissement de la masse d’air considérée, transportée à partir des zones urbaines ou industrielles fortement émettrices en espèces primaires. Par ailleurs, dans les sites trafic, l’ O_3 étant détruite par certains polluants primaires il est généralement absent sur ces sites très pollués. Pour toutes ces raisons, l’ O_3 est considéré comme le traceur majeur de la pollution photochimique.

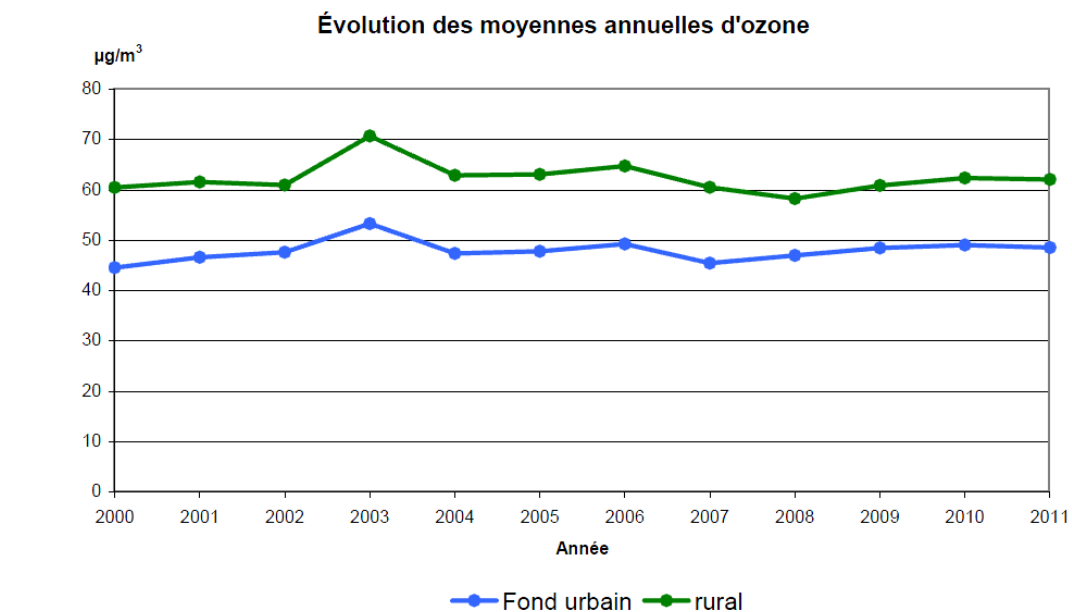


Figure 1-8 : Evolution de la concentration moyenne annuelle d'O3 de 2000 à 2011.

Tous les polluants cités précédemment présentent des concentrations variables dans l’atmosphère et contribuent d’une façon directe ou indirecte à la pollution atmosphérique. Le tableau 1-1 rappelle les principales sources de ces polluants, leurs effets sur la santé humaine

ainsi que les tonnages émis en France entre 2000 et 2009. Les Directives imposées par l'Union Européenne en matière de qualité de l'air associées aux normes européennes visant à la mise en application des directives et traduites en lois ou normes par l'état français (ou les états membres de l'UE) (Ministère du développement durable, 2010) ont permis de diminuer de près d'un facteur 2 les émissions des polluants primaires en environ 9 années.

Tableau 1-1 : Principaux polluants atmosphériques mesurés, leurs sources et leurs effets (El Yamani, 2006) avec les émissions en France reportées entre 2000 et 2009 (source : CEIP-WebDab reporté U-FR11 (2011-02-14)). 1 Gg = 10⁹ g

Principaux polluants atmosphériques	Sources principales	Effets sur la santé	Emissions enregistrée durant (Gg)	
			2000	2009
SO₂	Industries (installations thermiques, raffineries, etc.)	Très soluble, rapidement absorbé par les surfaces humides de la bouche, du nez. Irritant respiratoire, contribue à l'exacerbation des troubles bronchiques	632	302
NO_x (NO et NO ₂)	Industries de combustion (transports, installations thermiques, etc.)	Irritations de l'appareil respiratoire, crises d'asthme et bronchiolite	1575	1116
CO	Installations de combustion, transports, chauffage domestique	Atteinte du système nerveux central et des organes sensoriels.	6709	3950
O₃	Polluant secondaire formé suite à des réactions chimiques entre COV et NO _x en présence de rayons ultraviolets	Polluant irritant des yeux, de la gorge, et des poumons, peut altérer la fonction respiratoire et la résistance aux infections	-----	-----
COV-NM	Combustion, usage et évaporation des solvants et carburants industriels, etc.	Irritant des yeux et des poumons peuvent engendrer des bronchites par intoxication chronique. A long terme, ils seraient responsables de cancers	1706	878
PM₁₀ et PM_{2.5}	Véhicules (en particulier diesel) et combustion de certaines industries	Irritant très actif par altération des fonctions respiratoires, crises d'asthme avec à long terme des bronchites chroniques	PM10 : 501 PM2.5 : 368	PM10 : 364 PM2.5 : 251

1.2. Réglementation

En raison du caractère dangereux des différents polluants et la dégradation de la qualité de l'air provoquant une dégradation de la santé des populations, la surveillance de la qualité l'air est devenue une nécessité. Cette surveillance a débuté au début des années 60

pour ce qui concerne les dépôts des particules, puis elle a été remplacée, au début des années 70, par la surveillance de polluants gazeux, cette surveillance a été organisée par l'apparition des directives européennes de la gestion de la qualité de l'air dont l'Union Européenne définit le cadre. La stratégie de surveillance et de gestion de la qualité de l'air ambiant extérieur se base sur la Directive Européenne 96/62/CE (Mathé, et al., 2009). Ces Directives sont mises en place dans le dispositif national de surveillance de la qualité de l'air avec l'appui technique des laboratoires de référence nationaux comme par exemple le Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) composé de l'INERIS, du LNE, et de l'Ecole des Mines de Douai. Les polluants gazeux dont les émissions dans l'atmosphère sont surveillées doivent répondre aux directives imposées par l'Union Européenne relatives aux «polluants réglementés». Le tableau 1-2 donne la liste des Directives les plus importantes imposées par l'Union Européenne pour la réglementation et la protection de la qualité de l'air. Au niveau national, la loi n°96-1236 du 30 décembre 1996 sur l'air et l'utilisation rationnelle de l'énergie (loi LAURE) énonce le principe du droit de chacun à respirer un air qui ne nuise pas à sa santé.

Tableau 1-2 : Les Directives imposées par l'Union Européenne pour la protection et la réglementation de la qualité de l'air.

Directive	Date de parution	Cible
80/779/EEC (C.E., 1980)	15-juillet-1980	SO ₂ et PM ₁₀
82/884/EEC (C.E., 1982)	3-December-1982	NO _x
85/283/EEC (C.E., 1985)	7-Mars-1985	
92/72/EEC (C.E., 1992)	21-Septembre-1992	O ₃
1999/30/CE (C.E., 1999)	22-Avril-1999	SO ₂ , NO ₂ & NO _x , PM ₁₀ , Pb
2000/69/CE (C.E., 2000)	16-Novembre-2000	C ₆ H ₆ , CO
2002/3/CE (C.E., 2002)	12-Février-2002	O ₃
2004/107/CE (C.E., 2004)	15-décembre-2004	As, Cd Hg, Ni et HAP
2008/50/CE (C.E., 2008)	21-mai-2008	SO ₂ , NO ₂ & NO _x , PM ₁₀ , PM _{2,5} , Pb, C ₆ H ₆ , CO, O ₃ .

Il est à noter que la réglementation a évolué et qu'à ce jour, c'est la Directive 2008/50/CE qui est en vigueur et définit le cadre de surveillance d'un grand nombre de polluants sauf pour les métaux lourds et les HAP pour lesquels la directive 2004/107/CE est toujours en vigueur. Ces Directives imposées par l'Union Européenne introduisent des valeurs limites et des seuils d'information et d'alerte pour la protection de la santé humaine et de la végétation (Mathé, et al., 2009). La notion d'objectif de qualité a été introduite et vise à réduire année après année la présence des polluants dans l'air. Le tableau 1-3 répertorie les valeurs des objectifs de la qualité de l'air avec les valeurs limites et les seuils d'alerte et d'information en vigueur à ce jour au vu des Directives 2008/50/CE et 2004/107/CE.

Ces limites et valeurs imposées par les Directives Européennes sont renforcées par les termes de la loi n° 96-1236 du 30/12/96 (loi LAURE). Cette loi définit un sens plus large de la pollution atmosphérique en précisant (Mathé, et al., 2009) :

- les conditions de mise en place de dispositifs de surveillance de la qualité de l'air dans les agglomérations,
- les objectifs de la qualité de l'air,
- la prise en compte de mesures d'urgence en cas d'alerte à la pollution atmosphérique,
- l'information du public,
- et les mesures fiscales et financières favorisant des systèmes et des énergies peu polluantes.

Tableau 1-3 : Récapitulatif sur les objectifs de qualité et les limites imposées par la Loi n° 96-1236 parue le 30/12/96 sur l'air et l'utilisation rationnelle de l'énergie sur les polluants (mise à jour des valeurs : Février 2010 et directive 2008/50/CE). AOT40* : exposition accumulée à l'O₃ au-delà de seuil de 40 ppb.

Gaz Cible	Objectif de qualité	Valeurs limites pour la protection de la santé humaine	Valeur limite pour la protection de la végétation	Seuil d'information	Seuils d'alerte
NO₂	40 µg/m ³ en moyenne annuelle	200 µg/m ³ en moyenne horaire à ne pas dépasser plus de 18 heures par an (Centile 99,8)		200 µg/m ³ en moyenne horaire	400 µg/m ³ en moyenne horaire
		44 µg/m ³ (40 µg/m ³ en 2010) en moyenne annuelle			ou si 200 µg/m ³ en moyenne horaire à J-1 et à J, et prévision de 200 µg/m ³ à J+1
NO_x			30 µg/m ³ en moyenne annuelle		
O₃		à compter de 2010 120 µg/m ³ maximum journalier de la moyenne sur 8 heures à ne pas dépasser plus de 25 jours par an(en moyenne sur 3 ans)	6 000 µg/m ³ .h en AOT40, calculée à partir des valeurs sur 1 heure de mai à juillet (en moyenne sur 5 ans)	180 µg/m ³ en moyenne horaire	1er seuil : 240 µg/m ³ moyennes horaires pendant 3 heures consécutives
					2° seuil : 300 µg/m ³ moyennes horaires pendant 3 heures consécutives
					3° seuil : 360 µg/m ³ en moyenne horaire
CO		10 mg/m ³ soit 10 000 µg/m ³ pour le maximum journalier de la moyenne glissante sur 8 heures			
SO₂	50 µg/m ³ en moyenne annuelle	350 µg/m ³ en moyenne horaire à ne pas dépasser plus de 24 heures par an (Centile 99,7)	20 µg/m ³ en moyenne annuelle et en moyenne sur la période du 1er octobre au 31 mars	300 µg/m ³ en moyenne horaire	500 µg/m ³ en moyenne horaire pendant 3 heures consécutives
		125 µg/m ³ en moyenne journalière à ne pas dépasser plus de 3 jours par an (Centile 99,2)			
BENZÈNE	2 µg/m ³ en moyenne annuelle	5 µg/m ³ en moyenne annuelle			
BENZO-(a)-PYRÈNE	à compter de 2013 1ng/m ³ en moyenne annuelle dans la fraction PM ₁₀				
PM₁₀	30 µg/m ³ en moyenne annuelle	50 µg/m ³ en moyenne journalière à ne pas dépasser plus de 35 jours par an (Centile 90.4)		80 µg/m ³ en moyenne sur 24 heures	125 µg/m ³ en moyenne sur 24 heures
		40 µg/m ³ en moyenne annuelle			

2. Mesures et méthodes normalisées

La Directive 2008/5/CE et l'AFNOR¹¹ ont préconisé des méthodes spécifiques de mesure pour chaque polluant. Ces méthodes normalisées permettent d'accéder aux niveaux de concentration des gaz ciblés avec une grande précision correspondant aux objectifs de qualité spécifiques dans la directive. Le tableau 1-4 présente les différentes méthodes de mesure spécifiques à chaque polluant avec la référence de la norme associée.

Tableau 1-4 : Les différentes méthodes normalisées pour les gaz réglementés (Houdret, et al., 2008).

Gaz	Méthode de mesure	Norme
NO _x	chimiluminescence	EN 14211 (2005)
CO	rayonnement infrarouge non dispersif	EN 14626 (2005)
O ₃	photométrie UV	EN 14625 (2005)
SO ₂	fluorescence UV	EN 14212 (2005)
Benzène	prélèvement actif et chromatographie	EN 14662-1 (2005) EN 14662-2 EN 14662-3

Des analyseurs automatiques d'air ont été développés spécifiquement pour effectuer des mesures selon les méthodes normalisées. Ils peuvent être utilisés dans des stations de mesures de la qualité de l'air pour assurer le suivi au quotidien de la pollution atmosphérique (Sivacoumar, et al., 2001).

2.1. Différentes méthodes de mesures spécifiques aux polluants

Les méthodes normalisées listées dans le tableau 1-4 pour la mesure des gaz réglementés nécessitent l'utilisation d'analyseurs intégrant des composants complexes pour assurer des mesures répondant aux exigences de qualité requises dans les directives et normes (source de radiation UV ou IR, filtre, four...) (Toupance, et al., 2004) (Skoog, et al., 1992) ce qui rend ces analyseurs assez complexes et coûteux. Le tableau 1-5 répertorie les principales méthodes de mesures utilisées par ces analyseurs en précisant leurs atouts et leurs faiblesses.

¹¹ Association Française de Normalisation.

Tableau 1-5 : Récapitulatif des méthodes d'analyses de gaz (Ngo, 2006).

Technique utilisée	Gaz détecté	Minimum détectable	Temps de réponse	Avantages	Inconvénients
Chromatographie en phase gazeuse	COV (selon détecteur et colonnes)	<<1ppm	5 à 40 min	Sensible, spécifique	Etalonnage nécessaire, peu portable, méthode destructive
Spectrométrie de masse	Tout composé sous forme gazeuse	ppb à ppm	≈ 1 min	Sélective, sensible, quantification	Matériel lourd, interprétation de spectres
Absorption infrarouge	CO, CO ₂ , SO ₂ , HCl, HC	≈ ppm	≈ 1 min	sélective, sensible	Interférence avec vapeur d'eau
Chimiluminescence	NO, NO ₂	1 ppb	≈ 1 min	Très spécifique	Maintenance et étalonnage délicat, méthode destructive
Fluorescence	SO ₂	<0.1 ppm	≈ 1 min	Peu d'interférence	Spécifique SO ₂

Ces méthodes assurent généralement des mesures sélectives avec une grande précision en un temps de réponse relativement court (Toupance, et al., 2004). Leurs points négatifs majeurs résident dans leur non portabilité et leur complexité d'utilisation. Ces outils ont besoin d'être étalonnés régulièrement et leur maintenance peut s'avérer coûteuse.

2.2. Réseaux de surveillance de la qualité d'air

Des réseaux de surveillance de la qualité de l'air et de mesures de la pollution atmosphérique sont répartis en France et dans toute l'Europe. On compte au minimum un réseau par région en France comme montré sur la figure 1-9.

Chaque réseau prend en charge la surveillance de l'air dans sa région, et met en place différentes stations de mesures, contenant chacune un ou plusieurs analyseurs physicochimiques d'air, basés sur les méthodes de mesures normalisées listées précédemment (c.f. tableau 1-4). Ces stations sont placées en différents points fixes des principales agglomérations pour répondre aux exigences requises par les Directives. Les mesures de la concentration de chaque polluant dans l'air sont relevées en continu pendant 24 heures lorsqu'il s'agit de mesures automatiques ou sont réalisées à une fréquence régulière.

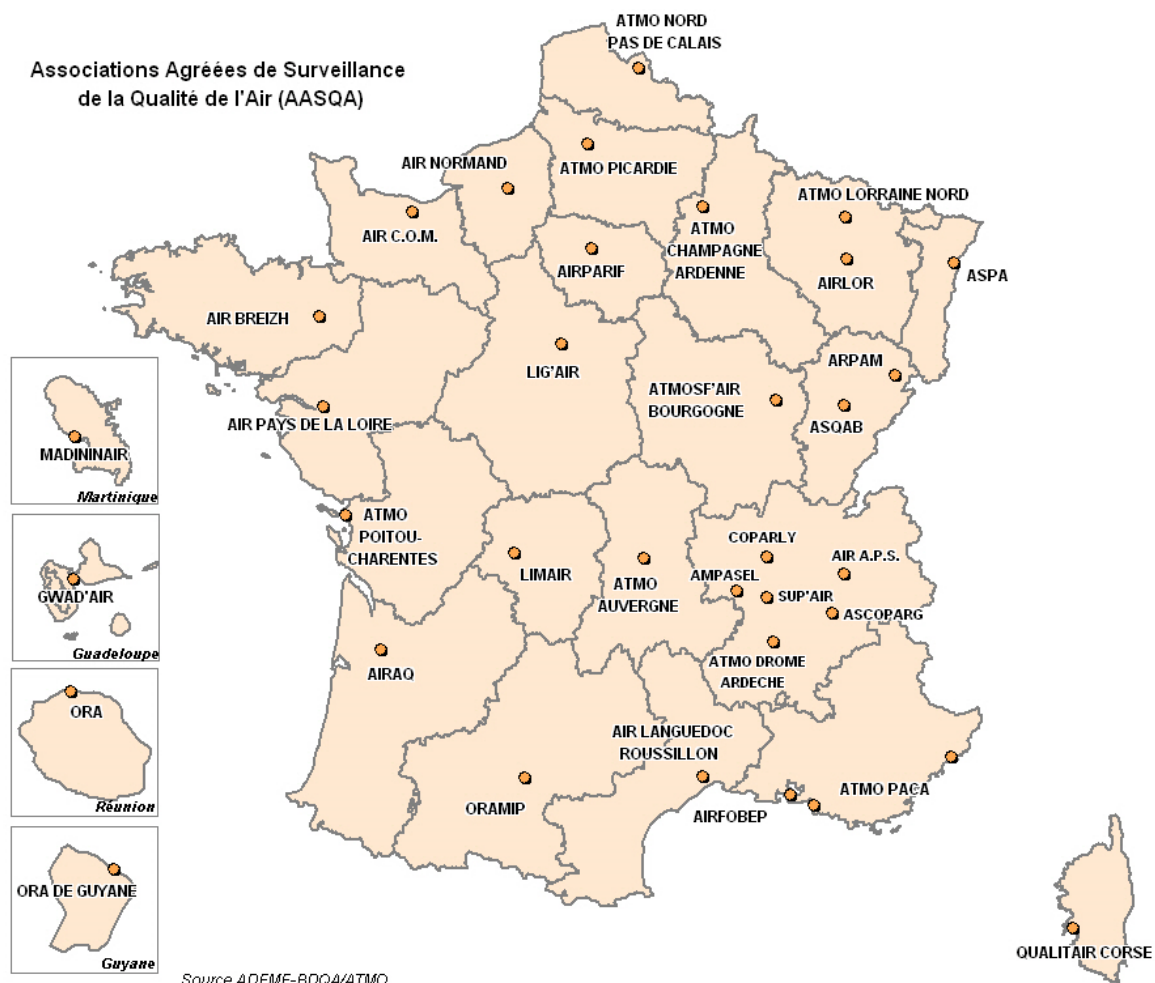


Figure 1-9 : Répartition des réseaux de mesure de qualité d'air en France¹².

Ces stations de mesures nécessitent une équipe de personnel qualifié (des techniciens, informaticiens, ingénieurs...) afin d'assurer le bon fonctionnement du matériel, son étalonnage, la récupération et le traitement des données. Ces équipes doivent avoir également des compétences en maintenance électrotechnique et électronique des appareils. Outre le coût en ressources humaines, le budget de construction et d'équipement d'une seule station dépasse plusieurs dizaines de milliers d'euros. Cela limite fortement le nombre d'implantations possible qui se réduit généralement aux exigences requises par la Directive Européenne ou pour la surveillance de points particuliers comme autour des sites industriels.

¹² <http://www.developpement-durable.gouv.fr>

2.3. Stations de mesure de la pollution

Plusieurs paramètres et critères spécifiques sont attribués aux stations de mesure de la pollution par les Directives Européennes (Mathé, 2010). La classification des stations correspond aux problèmes de pollution atmosphérique rencontrés en fonction de la zone géographique où elles sont implantées. Elles peuvent appartenir à des zones urbaines ou des zones rurales, les zones urbaines sont les unités bâties et telles que la somme des populations des communes qui les composent soit supérieure à 2000 habitants, les zones rurales sont les communes qui ne font pas partie de la zone urbaine (ADEME, 2002). L'implantation répond elle-même à des critères de classification bien spécifiques. Ces critères dépendent du nombre d'habitants dans une zone urbaine, de la position de la station, du trafic automobile quotidien, du rapport entre les différents polluants rencontrés, et enfin de la hauteur du point de prélèvement (Mathé, 2010). On distingue 2 groupes de stations : les stations de fond et les stations de proximité (Mathé, et al., 2009). Les stations de fond réalisent un suivi de l'exposition moyenne des personnes et de l'environnement à des phénomènes de pollution atmosphérique, c'est le cas des stations dites « urbaines » et « périurbaines ». Les stations de proximité telles que les stations dites « industrielles » et « trafic », fournissent des informations sur les niveaux maximaux des concentrations des polluants mesurés dans leurs zones représentatives, ainsi que les niveaux d'exposition de la population située à proximité des sources. A chaque type de station, est dédiée une liste de polluants spécifiques. Nous pouvons classer en France 7 différents types de stations appartenant à ces deux groupes (ADEME, 2002):

- Station urbaine
- Station périurbaine
- Station rurale régionale
- Station rurale nationale

- Station industrielle
- Station trafic
- Station d'observation spécifique

Le tableau 1-6 présente les objectifs d'implémentation de chaque type de station, ainsi que les polluants surveillés et de leur rayon de représentativité. Il est à noter que la représentativité des stations de proximité est bien plus réduite que celles des stations de fond. Par ailleurs, les Directives Européennes 2000/69/CE et 1999/30/CE précisent que les stations devraient être de préférence représentatives d'une aire de plus de 200 m² pour les sites trafics, de plusieurs km² pour les sites urbains, et d'au moins 1000 km² pour les sites surveillant les écosystèmes (ADEME, 2002). Nous comptons en France 249 stations urbaines, 113 stations périurbaines et 87 stations trafics ainsi que 37 stations rurales régionales, 8 stations rurales nationales et 34 stations d'observation spécifique (ADEME, 2002).

Parmi les différents critères permettant de classer les types de station, on peut noter les suivants :

- Un rapport de la moyenne annuelle en concentration « R » de monoxyde d'azote sur celle de dioxyde d'azote constitue le marqueur pour la définition d'un site d'accueil d'une station trafic (Mathé, 2010) :

$$R = \frac{[NO]}{[NO_2]}$$

- Pour une valeur de R inférieure à 1,5 la station est classée comme urbaine ou périurbaine, pour une valeur de R supérieure à 1,5 on classe la station comme trafic.

Tableau 1-6: différents types des stations de surveillance de qualité de l'air (Mathé, et al., 2009) (Mathé, 2010) (ADEME, 2002).

Groupe de station	Type de station	Rayon de l'aire de représentation	Objectifs	Polluants mesurés	Nombre de station en 2002
Station de fond réalisant un suivi de l'exposition moyenne des personnes et de l'environnement.	urbaine	De 100 m à 2 km	Suivi de l'exposition moyenne de la population aux phénomènes de pollution atmosphériques dits de « fond » dans les centres urbains.	NO _x , O ₃ , SO ₂ , PM ₁₀ , COV	249
	Périurbaine	De 1 km à 5 km	Suivi de la pollution photochimique (notamment O ₃ et ses précurseurs) et éventuellement les polluants primaires. Suivi du niveau d'exposition moyen de la population aux phénomènes de pollution atmosphérique dits de « fond » à la périphérie du centre urbain.	NO _x , O ₃	113
	Régionale rurale	De 5 km à 25 km	Surveillance de l'exposition de la population et les écosystèmes à la pollution atmosphérique de « fond » (notamment photochimique) à l'échelle régionale. Participation à la surveillance de la qualité de l'air sur l'ensemble du territoire, notamment dans les zones rurales.	NO _x , O ₃ , NH ₃	37
	Nationale rurale	De 30 km à 500 km	Surveillance dans les zones rurales de la pollution atmosphérique dite de « fond » issue des transports de masses d'air à longue distance, notamment transfrontaliers.	NO _x , O ₃ ,	887
Station de proximité	Industrielle	De 100 m à 5 km	Source d'informations représentatives du niveau maximum auquel la population riveraine d'une source fixe est susceptible d'être exposée.	SO ₂ , COV, HAP, NO _x , HP, O ₃ *, CO*	163
	Trafic	De 10 m à 50 m	Source d'informations représentatives du niveau maximum auquel la population située en proximité d'une infrastructure routière est susceptible d'être exposée	NO _x , CO, SO ₂ *, Pb*, HAP*	87
Observation spécifique		Selon les besoins	Réponse à des besoins spécifiques tels que l'aide à la modélisation, la prévision...	Selon les besoins	34

*polluants optionnels

Pour l'implantation des stations trafic, la distance maximale à une voie automobile fréquentée ne doit pas dépasser 10 m pour être au plus près de la source. Les stations urbaine et périurbaine sont généralement sous l'influence d'une pollution photochimique forte en période estivale. Par contre, au niveau des sites trafic, les fortes concentrations de NO dans l'air provenant de la circulation automobile peut réagir avec l'O₃ présent dans l'air en le transformant en NO₂ ce qui constitue donc un puit de NO. Par conséquent, même en période de pollution photochimique, les sites trafic peuvent avoir une concentration très basse en O₃. Mesurer l'ozone en station trafic n'est pas donc pertinent (Mathé, 2010) et aucune des stations de type trafic ne dispose d'un analyseur en O₃ créant ainsi au niveau des jeux de données disponibles une différence significative.

2.4. Information du public – Indice ATMO

Plusieurs approches ont été choisies par les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA) afin de présenter d'une manière synthétique les informations caractérisant la pollution atmosphérique. La méthode la plus utilisée est l'indice ATMO. Cet indice varie de 1 (très bon) à 10 (très mauvais) selon la sévérité de la pollution. Un arrêté ministériel du 22 juillet 2004 réajusté par l'arrêté du 21/12/2011 afin de s'adapter à l'évolution des normes réglementaires, définit les échelles permettant de calculer l'indice ATMO dans les agglomérations comptant plus de 100000 habitants.

Cet indice est construit à partir de 4 sous-indices eux-mêmes calculés à partir des concentrations en SO₂, O₃, NO₂ et PM₁₀. L'indice final correspond à la valeur la plus élevée de ces 4 sous-indices¹³. Les données de base pour le calcul de chaque sous-indice sont :

- La moyenne des concentrations maximales horaires observées pour SO₂, NO₂ et O₃.
- La moyenne des concentrations journalières observées pour les PM₁₀.

¹³ www.airaq.asso.fr

Le tableau 1-7 donne le tableau de calcul des sous-indices. A noter que cette échelle est quasiment linéaire jusqu'à l'indice 7. L'indice 10 correspond aux seuils d'alerte des polluants.

Tableau 1-7 : Indice ATMO provenant de l'arrêté du 21 décembre 2011 modifiant l'arrêté 22 juillet 2004 relatif aux indices de la qualité de l'air.

Indice	Qualité	SO ₂ (µg/m ³)	O ₃ (µg/m ³)	NO ₂ (µg/m ³)	PM ₁₀ (µg/m ³)
1	Très bon	0-39	0-29	0-29	0-6
2	Très bon	40-79	30-54	30-54	7-13
3	Bon	80-119	55-79	55-84	14-20
4	Bon	120-159	80-104	85-109	21-27
5	Moyen	160-199	105-129	110-134	38-34
6	Médiocre	200-249	130-149	135-164	35-41
7	Médiocre	250-299	150-179	165-199	42-49
8	Mauvais	300-399	180-209	200-274	50-64
9	Mauvais	400-499	210-239	275-399	65-79
10	Mauvais	≥ 500	≥ 240	≥ 400	≥ 80

2.5. Conclusion

En conclusion sur cette partie, il apparaît que même si les mesures de ces stations de mesure de la qualité de l'air sont fiables et précises, elles comportent quelques inconvénients majeurs:

Manque de mobilité : les mesures sont toujours effectuées au même point géographique, ce qui rend les mesures dynamiques et/ou le maillage d'un site difficile, voire impossible (Johnsons, et al., 1973) (Hunter, et al., 1992).

Coût d'installation et de fonctionnement élevé : au coût associé à l'achat des différents analyseurs spécifiques, il convient d'ajouter le coût associé à la maintenance, ainsi que le besoin de personnels qualifiés pour répondre aux questions de complexité et difficultés de mise en œuvre des appareils.

Lorsqu'une évaluation de la qualité de l'air est souhaitée sans que cette évaluation corresponde à une mesure réglementaire dans la cadre de la mesure de la qualité de l'air, des méthodes complémentaires, simples, mobiles, et peu coûteuses peuvent être envisagées quitte

à perdre un peu de précision des mesures. Ces méthodes ne cherchent pas à remplacer les stations de mesures de la qualité de l'air, mais simplement à offrir des méthodes complémentaires plus simples pour répondre à des problématiques où la précision de mesures n'est pas un aspect prioritaire. Ainsi ces dernières années, le domaine des capteurs de gaz a connu un intérêt tout particulier dans le domaine de la qualité d'air.

3. Les capteurs de gaz

Les méthodes normalisées de mesure de la pollution atmosphérique déjà citées nécessitent des instruments de mesure très sophistiqués et complexes, ce qui les rend généralement surdimensionnés dans des applications où la précision des mesures et la détermination sélective de toutes les espèces gazeuses présentes dans une atmosphère polluée n'est pas une priorité ; c'est le cas du suivi qualitatif de la qualité de l'air. Dans cette optique, les recherches relativement récentes en matière de métrologie de la qualité de l'air se sont intéressées à l'utilisation de capteurs de gaz basés sur des concepts physiques ou physico-chimiques offrant des méthodes de mesure plus simples et une plus grande souplesse d'utilisation (Pijolat, 1997). Les progrès technologiques de la microélectronique et l'élaboration de nouveaux matériaux sensibles utilisés dans l'élaboration de ces capteurs de gaz (Lalauze, 1997) sont à l'origine des avancées majeures dans ce domaine. Ils sont généralement constitués d'un matériau dont l'une ou plusieurs de ses caractéristiques physiques sera sensible à la nature de l'environnement gazeux auquel il est soumis. Une modification de l'environnement sera alors généralement traduite en une variation de signal électrique aux bornes du matériau. Dans cette optique, ces capteurs doivent être capables de présenter une sensibilité, une fiabilité et une sélectivité acceptables selon les exigences des applications visées. Ces paramètres sont primordiaux dans des dispositifs contenant plusieurs capteurs de gaz comme les réseaux des capteurs de gaz et les nez électroniques. Dans cette

partie, sont présentés les différents types des capteurs de gaz, en indiquant leurs avantages et inconvénients et en terminant sur leurs utilisations dans les réseaux de capteurs.

3.1. Les différents types de capteurs de gaz

Par définition, un capteur est un composant dont le fonctionnement est basé sur un phénomène physique, chimique ou biologique permettant de transformer une grandeur physico-chimique en un signal généralement électrique. On peut citer en exemple les capteurs physiques de température, de vitesse, d'accélération, de pression. Un capteur de gaz transforme la concentration d'une espèce gazeuse dans un environnement gazeux en une variation de tension, de courant ou d'impédance électrique (Permuy, et al., 1994).

Les capteurs de gaz chimiques, sont les capteurs les plus utilisés pour la reconnaissance des espèces gazeuses dans un environnement gazeux complexe. Ils sont constitués d'une couche sensible et d'un système de transduction (Permuy, et al., 1994). Cette couche réagit avec son environnement gazeux qui en retour modifie ses propriétés physiques comme montré dans la figure 1-10. Cette modification peut se présenter sous la forme d'une variation de conductance ($\Delta\sigma$), d'une variation de masse (Δm), ou d'une variation de l'intensité lumineuse dans le cas d'un matériau fluorescent ($\Delta I_{lum.}$). Le système de transduction va traduire cette variation de propriété physique en un signal électrique (ΔI), une variation de l'intensité lumineuse émise (ΔI_{lum}) ou une variation de fréquence (ΔF) (cas d'une microbalance à quartz). Parmi ces capteurs chimiques, on peut citer : les capteurs à électrolyte solide, les capteurs catalytiques, les capteurs résistifs, les capteurs capacitifs (Debliqy, 2010).

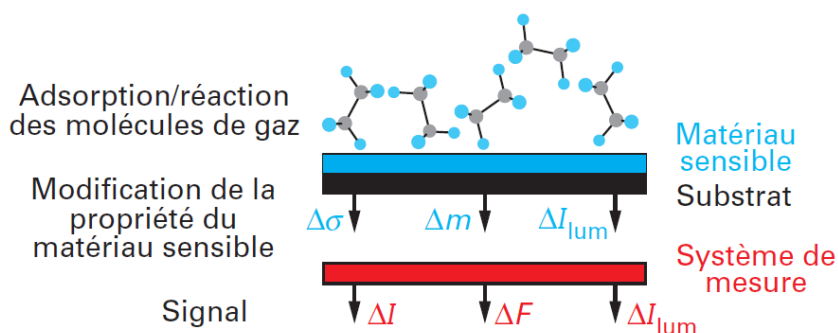


Figure 1-10 : la couche sensible dans un capteur chimique, et les molécules de gaz qui y interagissent (Thery-Merland, et al., 2006).

Ces capteurs de gaz chimiques sont classés en fonctions des différents types de phénomènes physiques sur lesquels ils se basent. La suite de ce chapitre donne un descriptif des types de capteurs de gaz chimiques les plus utilisés (Debliquy, 2010).

3.1.1. Les capteurs à électrolyte solide

Les capteurs à électrolyte solide, encore appelés « conducteurs ioniques » (Menil, 2008) sont des capteurs chimiques qui se basent sur les transferts de charges électriques entre les espèces gazeuses dans l'air et les ions mobiles dans l'électrolyte solide séparé par une membrane conductrice assurant le transport des ions. Le transport de ces ions induit une différence de potentiel entre les deux électrodes plongées dans l'électrolyte. Cette différence de potentiel est fonction de la concentration des espèces gazeuses en contact avec les ions (Jaffrezic-Renault, et al., 1994). Ces capteurs peuvent être de type potentiométrique ou ampérométrique. Les capteurs de gaz potentiométriques contiennent l'un de ces ions (O^{2-} , F^- , H^+ , Li^+ , Na^+ , Ag^+ , etc.). En revanche, les capteurs ampérométriques sont presque exclusivement basés sur des ions O^{2-} . Ils se basent sur la mesure de la différence entre la pression partielle d'oxygène dans l'électrolyte solide et l'oxygène de l'air, mais de nombreuses possibilités existent pour la détection d'autres gaz (CO , NO_2 , SO_2 , CO_2 , hydrocarbures...) (Menil, 2008). Ces capteurs nécessitent un système de chauffage pour porter l'électrolyte à une température de fonctionnement pouvant atteindre 800 °C. On donne

l'exemple d'une « sonde lambda » qui est un capteur potentiométrique d'oxygène dont l'électrolyte solide est à base de céramique de zircone stabilisée (voir figure 1-11). Ce type de capteur est actuellement monté sur les lignes d'échappement automobile pour surveiller les teneurs des gaz d'échappement (CO_2 , CO , NO_x ...) ce qui explique leur succès commercial (Menil, 2008). La sonde est munie de deux électrodes poreuses en platine, une au contact de l'air de référence et l'autre en contact direct avec le gaz d'échappement à travers une couche céramique poreuse. Ces deux électrodes sont séparées par un compartiment dans lequel la pression partielle d'oxygène est connue. L'électrolyte est chauffé à une température de l'ordre de 700-800 °C afin de permettre la conduction ionique. La différence de potentiel entre ces deux électrodes indique le taux d'échange des ions d'oxygène.

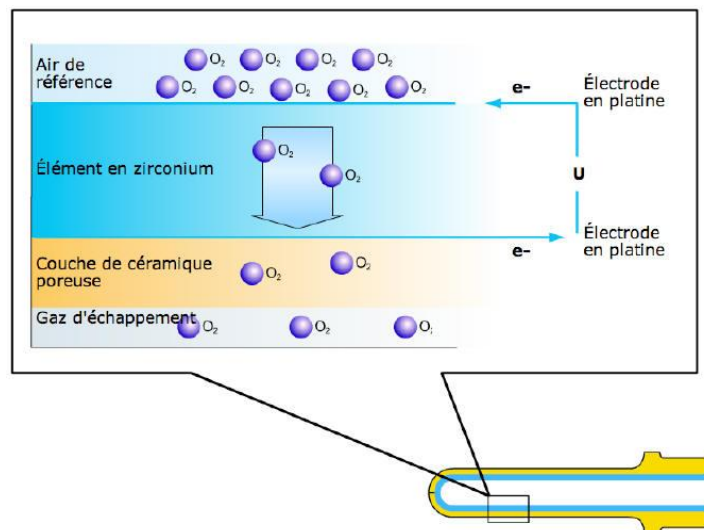


Figure 1-11 : Capteur à base d'électrolyte solide dans l'échappement des automobiles, sonde lambda (www.eobd.fr)

Ce capteur est également utilisé dans la régulation des conditions de combustion des moteurs automobiles afin d'aider à mieux contrôler les émissions de gaz toxiques tels que les hydrocarbures, les NO_x , et le CO (Ash, 2006).

Ces capteurs possèdent plusieurs avantages comme :

- la possibilité de miniaturisation,
- une bonne résistance aux chocs thermiques,
- une bonne résistance aux fortes pressions,

- une variation linéaire du signal avec la concentration d'oxygène.

Par contre, ils sont aussi connus pour leur sensibilité médiocre, leur complexité de mise en œuvre et d'implémentation dans des systèmes miniaturisés. Ils nécessitent également un étalonnage régulier, et enfin, ils ne permettent pas une mesure en continu. En raison de tous ces inconvénients, et tout particulièrement le dernier point cité concernant l'impossibilité d'un suivi en continu, nous avons écarté ce type de capteurs des options possibles pour la réalisation du réseau de capteurs que nous souhaitons mettre en œuvre.

3.1.2. Capteurs capacitifs

L'absorption d'une substance chimique dans un matériau modifie ses propriétés diélectriques. Ce mécanisme a conduit à la réalisation des capteurs chimiques capacitifs dans les années 80, où la capacité du capteur change au contact des espèces à détecter (Menil, 2008). Ce type de capteurs est largement utilisé pour la mesure de l'humidité relative. La figure 1-12 montre la composition d'un capteur capacitif pour la mesure de l'humidité relative.

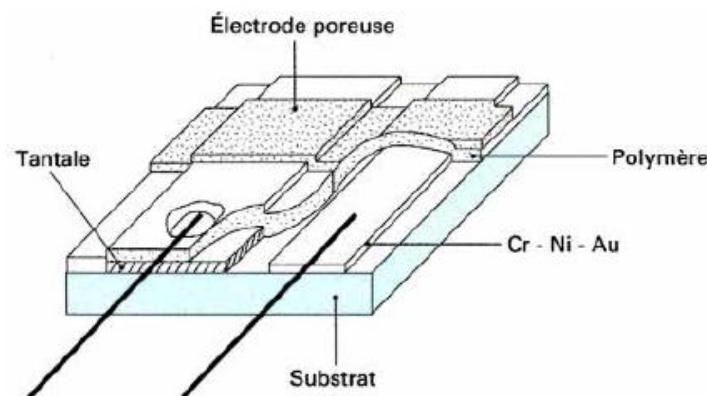


Figure 1-12 : Hygromètre capacitif à diélectrique polymère (Jaffrezic-Renault, et al., 1994)

Dans son principe, un hygromètre capacitif est constitué par une lame de polymère hygroscopique sur laquelle sont déposées deux électrodes métalliques poreuses, l'ensemble constituant un condensateur. Lorsque le polymère adsorbe les molécules d'eau, son volume

augmente et la distance entre les électrodes s'accroît, ce qui se traduit par une variation de capacité du condensateur (Debliqy, 2010) :

$$C = \frac{\epsilon_0 \epsilon_r (S)}{e}$$

Avec :

- ϵ_0 est la permittivité du vide ($= 8,852\text{pF/m}$)
- ϵ_r est la permittivité du matériau
- s est la surface des électrodes
- e est l'épaisseur du polymère.

Actuellement, différents types de ces capteurs sont développés afin de détecter des composés organiques comme l'éthanol ou le toluène (Menil, 2008). Ces capteurs présentent plusieurs avantages comme :

- une simplicité de mise en œuvre,
- une faible consommation d'énergie,
- un temps de réponse rapide,
- une intégration facile dans des applications portables ou embarqués,
- des possibilités de miniaturisation,

En revanche, leur sensibilité à l'humidité porte préjudice à la détection d'autres types de gaz.

Ce manque de sélectivité induit une perte de précision sur les mesures.

Compte-tenu de tous ces éléments lors de la réalisation du réseau de capteurs, seule la mesure du taux d'humidité sera effectuée à l'aide d'un capteur de type capacitif.

3.1.3. Capteurs résistifs, capteurs à oxydes semi-conducteurs (SnO_2)

Ces capteurs sont caractérisés par une couche sensible à base d'un matériau semi-conducteur (souvent le dioxyde d'étain, SnO_2). Un matériau semi-conducteur possède une

bande interdite suffisamment petite, par rapport à un isolant, pour que les électrons de la bande de valence puissent facilement rejoindre la bande de conduction. La figure 1-13 présente le diagramme des bandes d'énergie dans les cas de matériaux isolants, semi-conducteurs et conducteurs. Pour un métal, la bande interdite n'existe pas, les électrons se déplacent ainsi librement dans le matériau. Dans le cas d'un isolant, la bande interdite est trop large pour que le passage des électrons entre bande de valence et bande de conduction soit possible. Par contre, pour un semi-conducteur, un potentiel électrique suffisant appliqué à ses bornes, permettra le passage de porteurs à travers la bande interdite, induisant un faible courant électrique. Ce phénomène est fortement dépendant de la température et des impuretés présentes dans le matériau ou à sa surface.

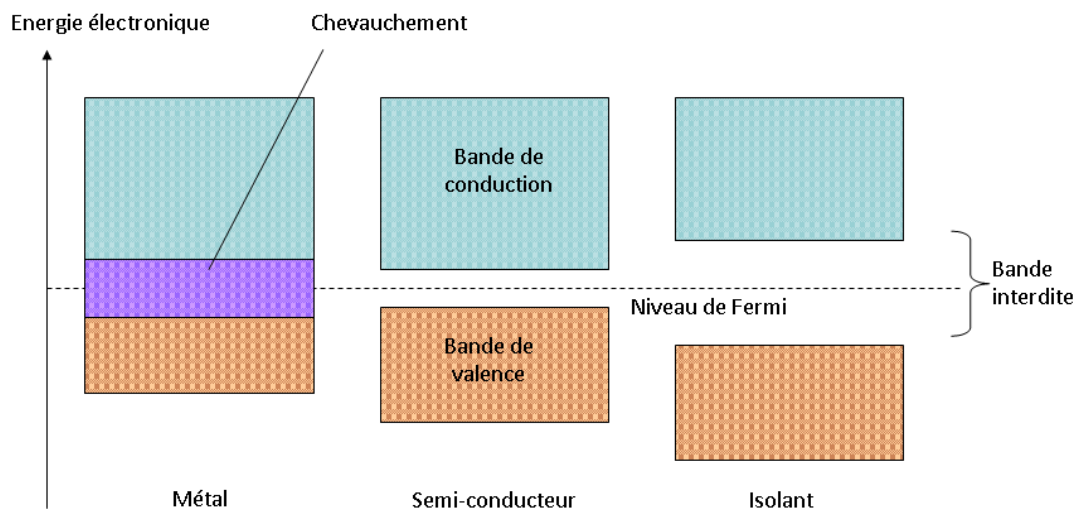


Figure 1-13 : Schéma théorique des bandes d'énergie en fonction de la nature d'un matériau.

Le principe de la détection dans l'air des capteurs de gaz à semi-conducteurs se base sur l'échange des électrons entre le matériau semi-conducteur polarisé de la couche sensible et des ions (O_2^-) provenant des molécules d' O_2 adsorbées à la surface. Pour former les ions O_2^- cette couche sensible doit être amenée à une haute température (entre 200° et 400°) (Debliqy, 2010). La variation de taux d'oxygène adsorbé à la surface change les caractéristiques du matériau ce qui provoquera un changement de résistance. Ainsi, la

diminution de la quantité d'oxygène adsorbée, soit par une diminution de la concentration de l'oxygène dans l'environnement du capteur, soit par une réaction à la surface du matériau, se traduit par une variation de conductivité (Jaffrezic-Renault, et al., 1994) (Debliquy, 2010). La figure 1-14 illustre le principe de ce fonctionnement. La présence de gaz oxydants ou réducteurs va occasionner des réactions d'oxydoréduction catalytiques à la surface mettant en jeu ces ions et donc bouleversant les équilibres et ainsi la conductivité électrique G . La variation d'énergie E de la barrière de potentiel suivant l'axe Z suite à l'adsorption d'oxygène sur la surface du matériel semi-conducteur induit la variation de conductance ΔG .

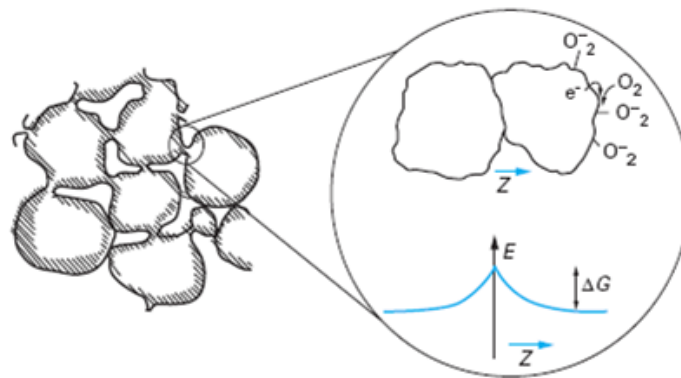


Figure 1-14 : représentation de la barrière de potentiel au joint de grains suite à l'adsorption d'oxygène (Debliquy, 2010).

Il existe plusieurs types de capteur à semi-conducteur qui diffèrent entre eux par la nature du semi-conducteur et par les types de gaz à détecter. Nous montrons dans le tableau 1-8 des exemples de différents types de matériaux et leur gaz cibles.

Tableau 1-8 : Famille de gaz détecté en fonction des différents types de semi-conducteur (Debliquy, 2010).

Semi-conducteur	Gaz détecté
SnO_2	Gaz combustibles (Méthane, propane, $\text{CO}\dots$)
WO_3	Ozone, NO_2
$\text{Cr}_{1.8}\text{Ti}_{0.2}\text{O}_3$	Ethanol et ammoniac

Pour assurer une cinétique d'adsorption et surtout de désorption nécessaire à leur fonctionnement, ces capteurs ne fonctionnent qu'à une température de quelques centaines de °C et sont donc munis d'une résistance de chauffage pour amener cette couche sensible à une

température précise (Kamionka, 2005). La relation entre la conductivité d'un semi-conducteur et la température est une relation exponentielle (Permuy, et al., 1994) (Nakata, et al., 1998). La figure 1-15 présente l'effet de la variation de la température de chauffage d'un capteur à semi-conducteur à base de SnO_2 par rapport à la détection de certains gaz. Nous remarquons que la température pour laquelle la sensibilité est optimale varie d'un gaz à l'autre. Le choix de la température de chauffage est donc très important suivant l'application et le gaz que l'on souhaite détecter (Kamionka, 2005).

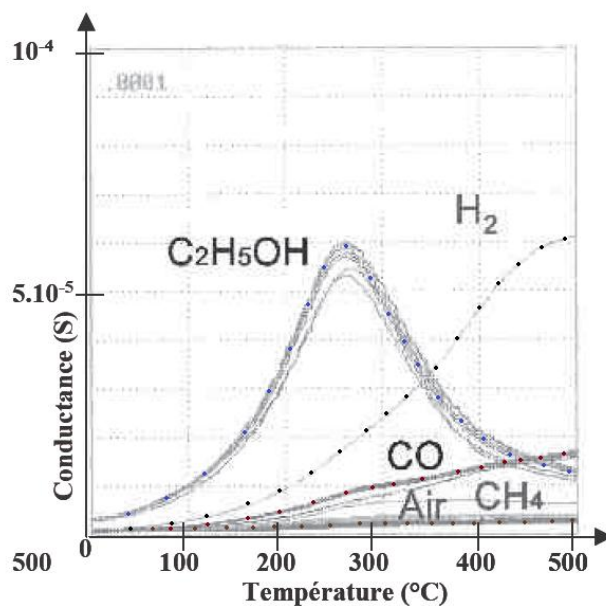


Figure 1-15 : Variation de conductance du SnO_2 pour des gaz spécifiés par rapport à la température (Kamionka, 2005).

Facile à concevoir, et de faible coût, ces capteurs sont largement utilisés pour des applications simples de suivi des nuisances dans l'air. Par ailleurs, comme ils consomment relativement peu d'énergie électrique, de l'ordre de quelques dizaines de mW, ces capteurs sont idéaux pour des applications visant une basse consommation électrique. La figure 1-16 illustre la composition d'un capteur semi-conducteur commercialisé, la couche sensible est déposée directement sur une résistance chauffante pour l'amener à sa température optimale. Ces capteurs ont été largement optimisés depuis une trentaine d'années, ce qui leur confère

une grande sensibilité pour la détection de certains gaz, un temps de réponse rapide qui peut atteindre quelques secondes, une faible taille, un faible coût (quelques euros) par rapport aux autres types de capteurs et une grande facilité d'utilisation dans des systèmes électroniques (Jaffrezic-Renault, et al., 1994).

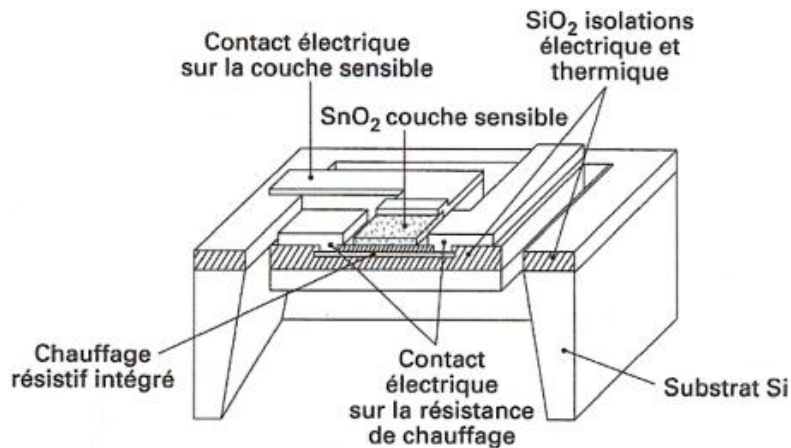


Figure 1-16 : Composition d'un capteur à semi-conducteur (Jaffrezic-Renault, et al., 1994)

Ces capteurs présentent toutefois quelques inconvénients dont les principaux sont:

- Une faible sélectivité : un capteur donné sera davantage sensible à une famille de gaz plutôt qu'à un gaz particulier.
- Une durée de vie limitée de 18 à 24 mois,
- Une dépendance à la température et à l'humidité ambiante (Delpha, et al., 1999) (Fort, et al., 2011).
- Des problèmes de dérive de ligne de base.
- Une non-linéarité de la réponse de ces capteurs à la concentration du gaz cible
- Un temps de recouvrement souvent plus long que le temps de réponse mais qui reste de quelques secondes (voir figure 1-17).

La figure 1-17 montre la réponse typique d'un capteur semi-conducteur en fonction de la concentration d'un gaz, nous remarquons la non linéarité de la réponse d'un tel capteur, ainsi que la présence d'un effet mémoire après la disparition de la concentration de gaz dû à l'accumulation des particules sur la couche sensible. Ce qui explique la présence d'un temps de recouvrement plus long que le temps de réponse.

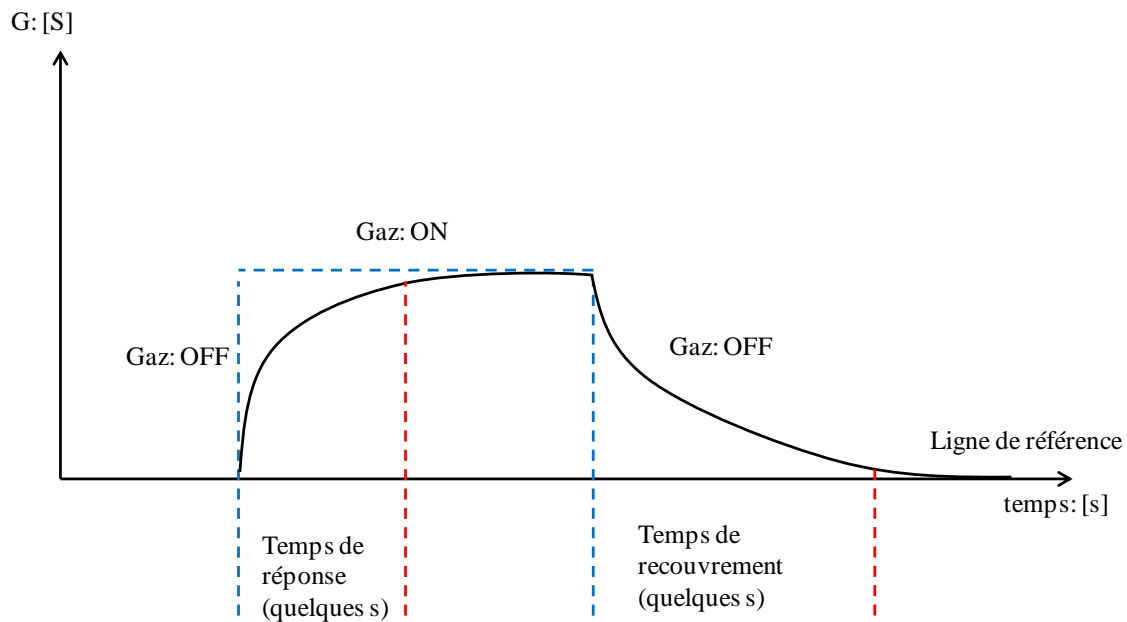


Figure 1-17 : réponse d'un capteur semi-conducteur à un échelon de gaz.

Même avec la présence de ces quelques aspects négatifs, la facilité d'utilisation, la bonne sensibilité et le temps de réponse rapide de ce capteur en font un bon candidat à notre application.

3.1.4. Autres types de capteurs - capteurs physiques

Pour détecter un gaz, il existe également des capteurs dit « physiques » (Menil, 2008) tels que les capteurs à absorption optique, les capteurs à photo ionisation (PID) et les capteurs à conductivité thermique. Ces capteurs « physiques » possèdent une durée de vie supérieure à celle des capteurs chimiques, mais en revanche ils présentent plusieurs inconvénients majeurs, rédhibitoires pour l'application visée dans le cadre de ce travail. Le principal inconvénient réside dans leur manque drastique de sensibilité, avec des limites de détection parfois supérieures à 1000 ppm. Par ailleurs, leurs conditions d'utilisation sont parfois délicates et ils sont souvent très peu sélectifs. Ces capteurs physiques ont donc été très vite écartés de notre choix pour la mise en place dans un réseau de caractérisation de la pollution atmosphérique, c'est pourquoi nous ne développons pas plus ici leur principe de fonctionnement.

3.2. Utilisation des capteurs dans la qualité de l'air

Comme chaque type de capteurs de gaz présente aussi bien des avantages que des inconvénients, il s'agira de trouver le meilleur compromis répondant aux contraintes imposées par le cadre du suivi de la qualité de l'air. Les capteurs idéaux doivent combiner facilité d'utilisation, bas coût, et temps de réponse rapide, tout en offrant une sensibilité et une sélectivité acceptables. Le tableau 1-9 répertorie les différents types de capteurs de gaz les plus utilisés, en rappelant brièvement leur principe de détection, leurs points forts et leurs faiblesses.

Dans le cadre du suivi de la pollution de l'air et de la construction d'un système portable à base de réseaux de capteurs, nous cherchons des capteurs ayant une bonne sensibilité avec un temps de réponse rapide, miniatures et faciles à utiliser et à implémenter dans des systèmes multi-capteurs, et de faibles coût et consommation électrique. Une lecture approfondie de ce tableau montre que les capteurs de gaz à semi-conducteurs (s-c) combinent l'ensemble des performances ciblées, et présentent surtout les limites de détection les plus basses. Par conséquent, ils constituent le meilleur choix pour une utilisation dans un système multi-capteurs pour le suivi de la pollution atmosphérique.

3.3. Caractéristiques métrologiques d'un capteur de gaz à s-c

Comme nous avons décidé de travailler avec des capteurs de gaz à s-c, il est important de rappeler les caractéristiques principales de ces composants. Les 3 paramètres les plus importants constituent ce qui est couramment appelé les « 3S » (Pijolat, 1997) : sensibilité, sélectivité et stabilité.

Tableau 1-9 : Classification de quelques types de capteurs de gaz selon leur principe de détection.

Type de Capteur	Grandeur mesurée	Avantages	Inconvénients	Limite de détection
Electrolyte solide	Tension ou courant	<ul style="list-style-type: none"> - Miniaturisation - Résistance aux chocs thermiques et pression - Variation linéaire du signal - Temps de réponse faible 	<ul style="list-style-type: none"> - Etalonnage nécessaire - Complexité d'utilisation - Durée de vie limitée (2 ans) - coût élevé (~100€) 	10-20 ppm
Capacitif	Capacité	<ul style="list-style-type: none"> - Simplicité d'implémentation - Faible coût (<10€) - Faible consommation électrique - Miniaturisation - Temps de réponse rapide - Longue durée de vie 	<ul style="list-style-type: none"> - Système de mesure complexe - Forte interférence avec l'eau - Manque de sélectivité 	~10 ppb
Résistif – semi-conducteur	Résistance ; Conductance	<ul style="list-style-type: none"> - Miniature - Faible coût (< 20€) - Faible consommation électrique - Facilité d'utilisation - Grande sensibilité - Temps de réponse faible - Large gamme de gaz détectables 	<ul style="list-style-type: none"> - Faible sélectivité - Durée de vie limitée (~2 ans) - Dépendance à la température - Problèmes de dérive 	10 ppb à 1 ppm
PID	Charge électrique	<ul style="list-style-type: none"> - Longue durée de vie 	<ul style="list-style-type: none"> - Manque de sensibilité et de sélectivité - Difficulté d'utilisation - Dimensions - Gamme de gaz détectables limitée - coût élevé (>100€) 	~100 ppm
Optique	Pic d'absorption	<ul style="list-style-type: none"> - Longue durée de vie - Miniaturisation 	<ul style="list-style-type: none"> - Difficulté d'utilisation - Coût élevé - Gamme de gaz détectables limitée 	~100 ppm
Conductimétriques	Conductivité thermique	<ul style="list-style-type: none"> - Longue durée de vie 	<ul style="list-style-type: none"> - Difficulté de miniaturisation et d'utilisation - Difficulté d'utilisation, - Non sélectivité. - Coût élevé (>100€) 	

3.3.1. Sensibilité

Par définition, il s'agit du rapport entre la variation élémentaire de la grandeur de sortie fournie par le capteur et la variation élémentaire de la grandeur à mesurer. Pour un capteur de gaz de type semi-conducteur, la sensibilité sera donc la variation de la conductance de sa partie sensible par rapport à la variation de la concentration de l'espèce gazeuse à laquelle il sera dédié. La sensibilité est exprimée en Siemens (ou Ohm^{-1}) par ppm comme montré sur la figure 1-18 (Kamionka, 2005). Si la réponse du capteur est linéaire, la sensibilité sera alors constante ce qui est rarement le cas pour les capteurs de gaz (Kamionka, 2005). Elle sera donc définie pour une concentration donnée.

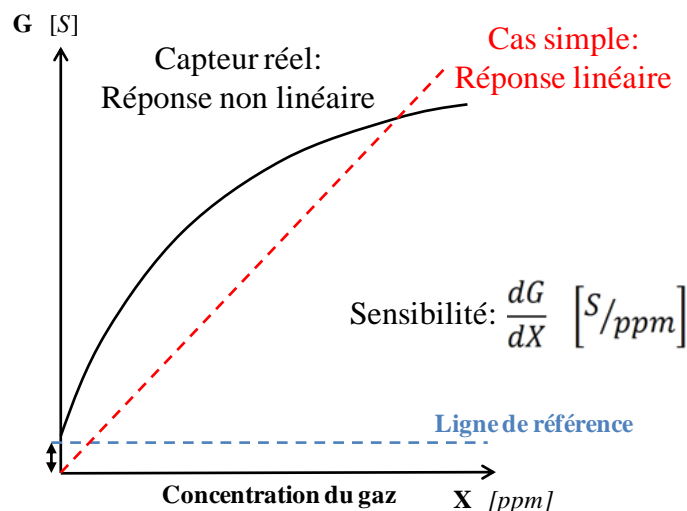


Figure 1-18 : réponse non linéaire d'un capteur à semi-conducteur

3.3.2. Sélectivité

C'est l'une des caractéristiques les plus importantes des capteurs de gaz. Par définition, la sélectivité est l'aptitude d'un dispositif à détecter quantitativement la présence d'un gaz indépendamment des variations de concentrations des autres gaz susceptibles d'interférer. On parle d'une sélectivité absolue lorsqu'un capteur ne répond qu'à une seule espèce gazeuse, et que sa réponse n'est pas affectée par la

présence d'une autre espèce. La sélectivité est dite relative si on dispose d'un capteur qui répond à différentes espèces susceptibles d'être présentes dans son environnement gazeux.

Les capteurs de gaz à semi-conducteurs sont généralement peu sélectifs, ce qui est une des grandes limitations de l'utilisation de ces capteurs (Pijolat, 1997). Cependant, en utilisant plusieurs capteurs mesurant le même environnement, et en exploitant l'ensemble de leurs signaux par des méthodes mathématiques adaptées, il est possible d'extraire de cette sélectivité partielle, des informations pertinentes. L'utilisation de plusieurs capteurs de gaz pour la détection des espèces gazeuses de l'air est considérée comme l'approche « multi-capteurs ». De la même façon, l'utilisation des méthodes mathématiques pour le traitement des signaux des capteurs est considérée comme l'approche « multi-variables » (Kamionka, 2005).

3.3.3. Stabilité

C'est un paramètre qui caractérise la dérive du signal des capteurs au cours d'une plus ou moins longue durée d'utilisation même si les conditions de mesures sont constantes.

Ces dérives peuvent être corrigées dans certains cas en effectuant périodiquement un calibrage « de zéro » (Tsujita, et al., 2005), c'est à dire une mesure dans une atmosphère de référence (généralement sous un air de composition contrôlée, appelé « air zéro »). Mais cette démarche possède deux inconvénients. D'une part il est nécessaire de disposer d'une atmosphère de référence, par exemple une bouteille de gaz étalon, ou un dispositif de « purification » de l'air. D'autre part, on suppose alors que seule la grandeur de référence $X(0)$ dérive, or souvent, c'est la sensibilité elle-même qui varie également. D'autres méthodes cherchent à modéliser cette dérive

(Romain, et al., 2010) (Tomic, et al., 2002), mais elles nécessitent de grosses ressources de calcul, et leur implémentation à un système temps réel s'ajoute à la complexité de l'outil de mesure (De Vito, et al., 2009). Les capteurs de gaz à semi-conducteurs sont particulièrement concernés par cette faiblesse (Menil, 2008) (Kamionka, 2005) (Korotcenkov, et al., 2011).

La facilité d'utilisation de ces capteurs avec leur faible coût rend ces capteurs assez populaires. Leur implémentation ainsi que leur utilisation dans différents systèmes de mesure des gaz est facile ce qui permet aussi la multiplication de ce type de capteurs dans un même système et permet donc l'obtention de la « signature » ou « empreinte » d'échantillons gazeux complexes, sans pouvoir en préciser la composition exacte (Menil, 2008). On parle alors de réseau de capteurs de gaz.

3.4. Réseau des capteurs de gaz - Nez électroniques

Des auteurs ont proposé de combiner l'utilisation de plusieurs capteurs pour la détection d'environnements gazeux complexes. Le premier essai expérimental utilisant plusieurs capteurs a été publié par Hartman et al en 1954 et 1964 en se basant sur plusieurs capteurs de type ampérométrique (Hartman, 1954) (Wilken, et al., 1964) (Gardner, et al., 1999). Un système multi-capteur combinant 8 de ces différentes cellules électrochimiques a été développé. Les réponses observées pour différents échantillons d'odeurs donnent différentes formes encore appelées « signatures ». Dans le laboratoire MICC à l'Ecole des Mines de Saint Etienne, l'utilisation des capteurs pour la détection des espèces gazeuses a débuté avec C.Pijolat durant sa thèse (Pijolat, 1986). Il a pu détecter la présence conjointe de benzène et d'éthanol en utilisant deux capteurs identiques à base de SnO_2 chauffés à des températures différentes. Ces essais ont ensuite été poursuivis par des études sur le développement des capteurs de gaz (Breuil, 1989) (Bruno, 1994) (Sauvan, 1999) (Montméat, 2002) (Rivière, 2004) (Perdreau, 2000) (Kamionka, 2005).

Les systèmes multi-capteurs de gaz ont été créés dans le but de détecter différentes espèces gazeuses spécifiques dans une atmosphère gazeuse artificielle (Gardner, et al., 1992) (Shurmer, et al., 1993) ou pour le suivi d'une pollution atmosphérique (Tsujita, et al., 2005). L'appellation de systèmes multi capteurs de gaz varie entre « réseaux de capteurs de gaz » et « nez électroniques » selon l'application ciblée. Les nez électroniques sont dédiés à la reconnaissance des odeurs, tandis que les réseaux de capteurs sont dédiés à la reconnaissance des espèces gazeuses au sens large, sans connotation olfactive. Il s'agit donc d'un système de capteurs à application spécifique (ASSS¹⁴) (Menil, 2008) composé de plusieurs capteurs chimiques possédant une sélectivité partielle, dite aussi relative. Insistons ici sur le fait que l'appellation « réseau de capteur » (différents capteurs implémentés dans un même ensemble) couramment utilisée, notamment dans ce manuscrit, peut prêter à confusion car il n'y a pas de répartition géographique des capteurs.

3.4.1. Approches de mesures quantitatives ou qualitatives

Si l'on dispose de plusieurs capteurs qui répondent à tous les gaz susceptibles d'être présents dans un environnement gazeux, avec des proportions différentes, un système de reconnaissance de formes peut alors être apte à discriminer plusieurs espèces ou groupe d'espèces gazeuses dans ce mélange. Ces systèmes peuvent donc être utilisés pour effectuer des mesures quantitatives dans des applications simples (Negri, et al., 2001), par contre pour des applications complexes comportant un grand nombre de gaz, ces systèmes seront limités à des mesures qualitatives ou semi-quantitatives (Helli, et al., 2004).

On distingue trois grands types d'applications suivant la finalité qualitative et/ou quantitative de l'analyse (Menil, 2008):

- **Sécurité et santé :** contrôle de qualité d'air intérieur, pollution, détection de gaz explosifs, toxiques ou dangereux. Dans ce cas on cherche à acquérir des

¹⁴ *Application Specific Sensor System, terme proposé par l'ETH zürich.*

informations les plus précises possibles. Il s'agit donc d'une approche quantitative (Tsujita, et al., 2005) (Zampolli, et al., 2004) (De Vito, et al., 2009) (Kamionak, et al., 2006).

- **Caractérisation d'une odeur ou d'un arôme** et toutes les recherches agro-alimentaires. Dans ce cas, on s'intéresse à la présence ou non d'une espèce ou d'un groupe d'espèces sans avoir besoin de connaître leurs concentrations. Il s'agit d'une approche purement qualitative (cas des nez électroniques) (Mielle, et al., 2001) (Capone, et al., 2000) (Ampuero, et al., 2003) (Cevoli, et al., 2011) (Lozano, et al., 2008) (Garcia, et al., 2006).
- **Applications hybrides**, pour lesquelles une évaluation des odeurs et plus largement d'espèces gazeuses sans quantification exacte est recherchée : usines, sites industriels, sites de compostages, bassins de décantation, épandages, eaux d'égout... (Nake, et al., 2005) (Romain, et al., 2005) (Romain, et al., 2000) (Bitter, et al., 2010). C'est le cas d'un nez électronique et il s'agit généralement d'une approche semi-quantitative (Haddi, et al., 2011).

3.4.2. Architecture d'un système multi-capteur

En général, les capteurs utilisés dans ces systèmes sont des capteurs à semi-conducteurs, soit d'origine commerciale (Negri, et al., 2001), soit réalisés par leurs concepteurs pour une application spécifique (Kamionka, 2005) (Althainz, et al., 1996) (Yamazoe, et al., 2009) (Sarry, et al., 1999). Ces systèmes sont exploités à l'aide de méthodes de reconnaissance de formes et d'analyse numérique (Chaiboun, et al., 2007). En utilisant un groupe de capteurs, dont la réponse individuelle est différente en fonction de la nature des gaz susceptibles d'être présents, on constitue une empreinte qui peut être unique et spécifique à chaque mélange gazeux, sous réserve d'appliquer un traitement des signaux adéquat. La recherche de cette

empreinte est l'approche multi-variable. Cet aspect est fortement exploité dans toutes les études sur les réseaux des capteurs et nez électronique (Tsujita, et al., 2005) (Thery-Merland, et al., 2006) (Getino, et al., 1999). La figure 1-19 présente, l'architecture de base d'un système multi-capteurs associée à un traitement de signal. Les signaux de plusieurs capteurs de gaz exposés ensembles au même environnement gazeux sont exploités afin d'obtenir une information adéquate sur les espèces gazeuses présentes.

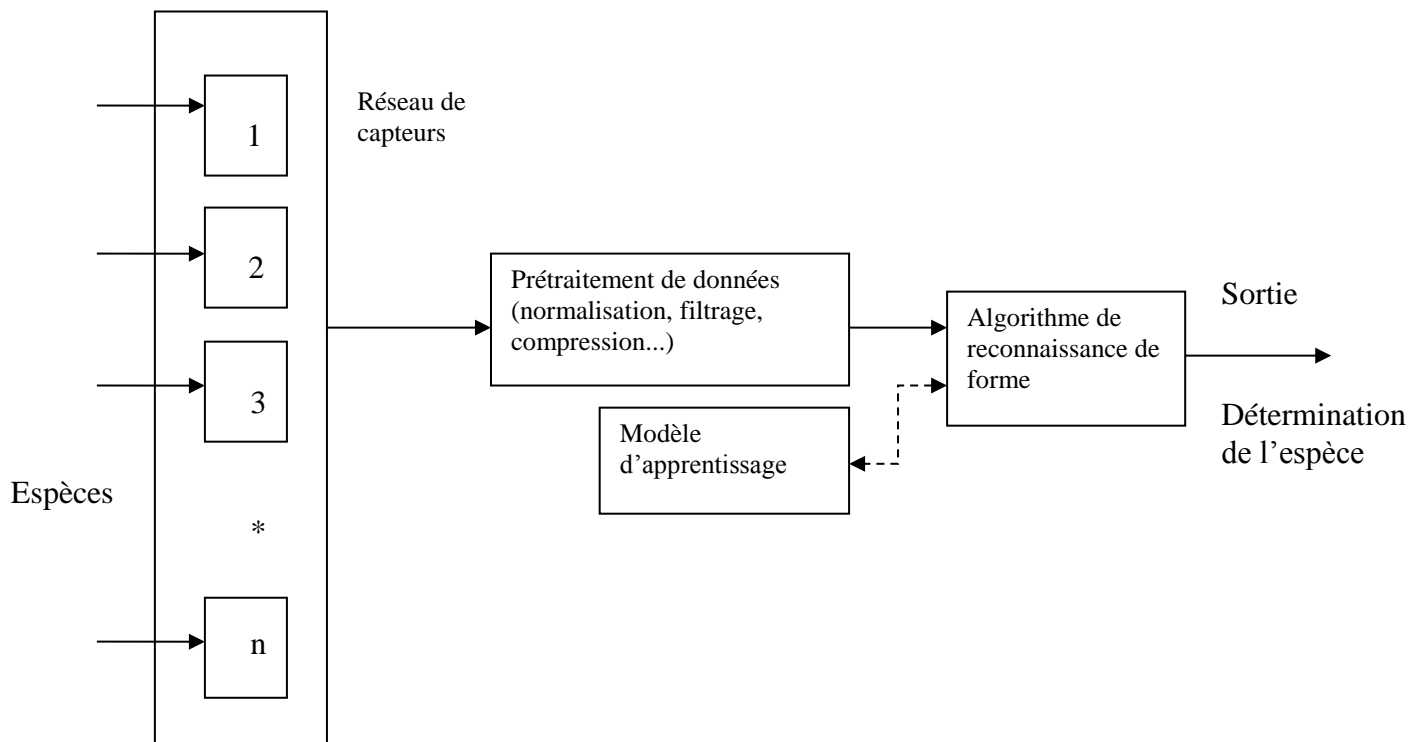


Figure 1-19 : Schéma de base de traitement de signal d'un réseau de capteurs

L'utilisation d'une approche multi-variable est une nécessité dans un réseau de capteurs non sélectifs (Shaffer, et al., 1999) (Sommer, et al., 1995). Les méthodes principalement utilisées comme les méthodes de reconnaissance de formes (Fu, et al., 2007) ont comme objectif de trouver une corrélation entre les signaux de sortie de ces capteurs et la composition chimique de leur environnement gazeux (Brudzewski, et al., 2007). Elles peuvent être divisées en deux groupes : les méthodes supervisées et les méthodes non supervisées. La technique supervisée consiste à repérer les espèces gazeuses par comparaison aux espèces

connues pendant l'étalonnage du système. Dans ce cas, une démarche d'étalonnage des capteurs sous différentes espèces gazeuses à mesurer sera nécessaire. Plus ces espèces sont nombreuses plus la démarche d'étalonnage sera longue et complexe. Cet étalonnage permettra de construire un modèle de comportement qui sera utilisé par un algorithme mathématique désigné. Ceci est décrit dans la partie désignée par « modèle d'apprentissage » présentée dans cette architecture sur la figure 16. Par contre, l'objectif de la technique non supervisée est de détecter et de discriminer les espèces inconnues en amplifiant les différences entre leurs vecteurs d'entrées associés (Gardner, et al., 1999). Ces méthodes, qui font aussi partie de l'analyse multi-variables, seront développées dans la section relative aux réseaux de neurones, mais par défaut, les méthodes que nous exposons sont supervisées.

4. Analyse multi-variable

L'analyse multi-variable est un outil mathématique utilisé en l'instrumentation dès les années 70 (World, 1995), en l'analyse chimique, pour interpréter les résultats des méthodes spectroscopiques et est souvent appelée pour cela « chimiométrie » (Breuil, 1997) (Otto, 2007). Aujourd'hui son utilisation touche presque tous les domaines scientifiques où la modélisation d'un nombre croissant de variables exige l'utilisation d'outils mathématiques adéquats. L'analyse multi-variable consiste à trouver des corrélations entre différentes données, ce qui est particulièrement adapté à la reconnaissance de signatures de la pollution atmosphérique issues des données fournies par les réseaux de capteurs.

4.1. Principes de modélisation multi-variables

La modélisation multi-variable consiste à modéliser les variations d'un certain nombre de variables dites *difficilement mesurables* appelées variables Y en fonctions d'autres variables dites *facilement mesurables* appelées variables X . Une modélisation mathématique suffisamment élaborée peut en théorie modéliser n'importe quel système réel complexe à

partir de données expérimentales (Munz, et al., 2009). Cependant, plus une méthode de modélisation est complexe, plus elle coûtera cher en terme de poids de calcul, et surtout moins elle aura de chances d'être extrapolable à d'autres données. De plus, la rapidité et la linéarité de la modélisation dépendent du type de méthodes choisis. En général, toutes les méthodes de modélisation possèdent des principes communs (Breuil, 1997).

Par exemple, la recherche par approches algébriques ou itératives d'un modèle mathématique cherche à établir la fonction vectorielle $Y=f(X)$ reproduisant au mieux des variables Y (sorties), par exemple des concentrations gazeuses, à partir de variables X (entrées), par exemple des signaux de capteurs. Dans ce cas, la fonction f est le prédicteur. Ce modèle mathématique est calculé lors d'une phase dite « étalonnage du modèle ». On distingue deux opérations résumées dans la figure suivante. Lors de l'étalonnage (ou modélisation), toutes les variables X et Y doivent être connues pour calculer le prédicteur. La phase dite de « prédiction », utilise le prédicteur et les X variables pour calculer les Y variables inconnues. Cette partie utilise le modèle pré établi comme montré dans la figure 1-20.

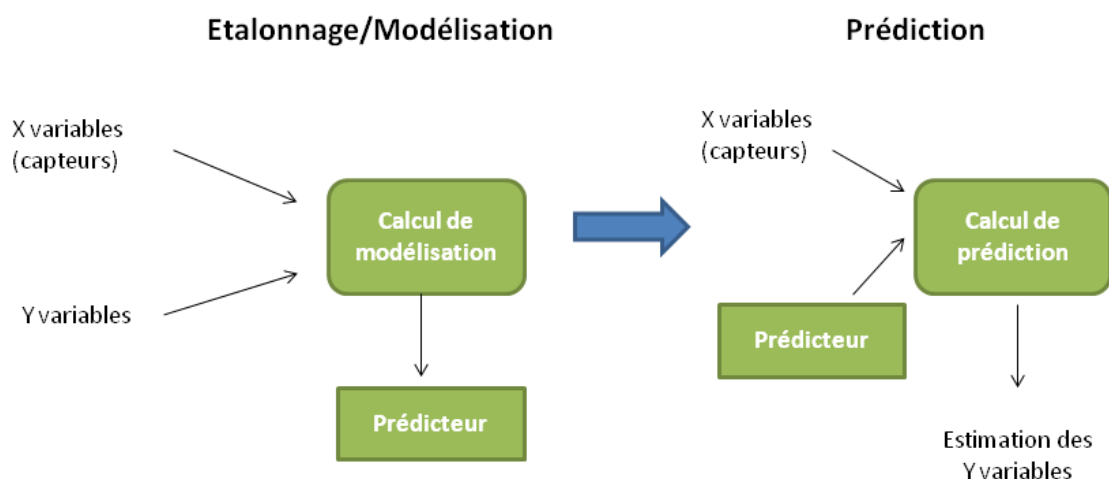


Figure 1-20 : principe d'analyse multi-variable

4.2. Problème de sur-modélisation :

Il est toujours possible en augmentant la complexité d'un modèle, d'aboutir théoriquement à un meilleur résultat pour un jeu de données d'étalonnage donné. Mais un « bon » modèle calculé à partir d'un groupe d'échantillons n'est pas forcément extrapolable à d'autres. Ainsi, si le modèle est trop précis, c'est à dire plus complexe que nécessaire, on risque de modéliser non seulement les phénomènes physiques, mais aussi les particularités des échantillons utilisés pour l'étalonnage, notamment l'erreur aléatoire, rendant dès lors ces modèles moins performants pour d'autres échantillons.

Par ailleurs, on peut montrer que, lors de la prédiction, même si le modèle est correct physiquement (erreur de modélisation négligeable), plus le modèle est complexe plus les erreurs aléatoires liées à la mesure sont statistiquement amplifiées lors de la prédiction (Kamionka, 2005). On définit ainsi le « taux d'amplification de l'erreur relative » α :

$$\frac{\Delta Y}{\|Y\|} = \alpha \cdot \frac{\Delta X}{\|X\|}$$

ΔX représente l'erreur aléatoire sur la X -variable X mesurée, et ΔY , appelée « erreur liée à la mesure », la conséquence de cette erreur sur la Y -variable calculée Y . La variable α représente la « sensibilité au bruit » du modèle, elle augmente lorsque la complexité du modèle augmente. Elle représente la propagation de l'erreur aléatoire à travers le modèle obtenu (Kamionka, 2005).

Ces problèmes, liées à une complexité trop grande du modèle sont dits de « surmodélisation ».

A l'inverse, si le modèle est trop simple par rapport à la réalité physique, on a alors une erreur dite « de modélisation », on parle de « sous-modélisation ».

Ainsi, pour obtenir le meilleur modèle, il faut trouver le modèle où l'erreur de prédiction soit minimale comme montré dans la figure 1-21, et donc un équilibre entre sous et sur-modélisation.

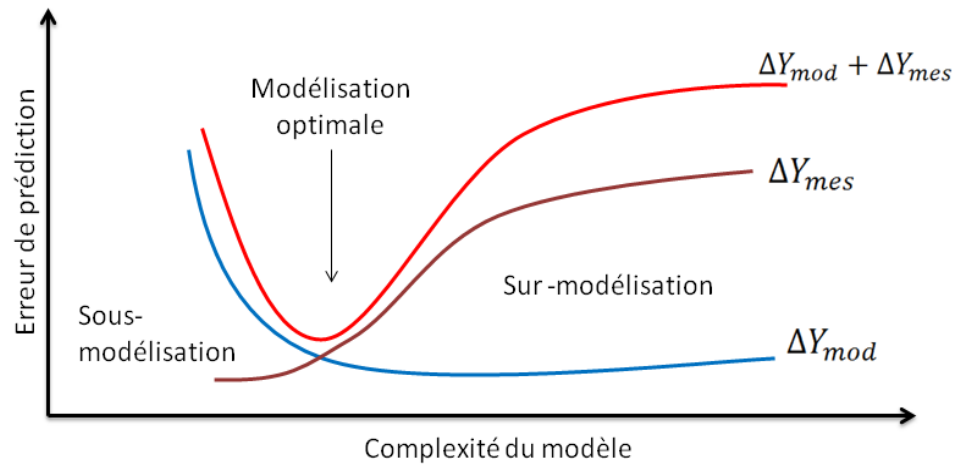


Figure 1-21 : erreur de prédiction en fonction de la complexité du modèle.

C'est la raison pour laquelle tout modèle doit être optimisé, testé et validé à l'aide de données distinctes des données d'étalonnage, cette étape sera appelée l'étape de validation.

4.3. Les différentes méthodes de modélisation

Une des principales difficultés sera donc le réglage adéquat de la complexité du modèle, la méthode la plus simple pour éviter cette sur-modélisation étant de tester tout modèle avec des échantillons n'ayant pas servi à le construire. Cet aspect sera développé lors de la présentation des algorithmes.

On peut classer les méthodes multi-variables en trois catégories : les méthodes linéaires simples, les méthodes linéaires complexes, et les réseaux de neurones (Gardner, et al., 1999).

4.3.1. Méthodes linéaires simples

Ce sont des méthodes linéaires peu puissantes mais capables de résoudre des problèmes simples basés sur des systèmes d'équations multi-variables linéaires. Ces méthodes, comme la

régression multilinéaire (MLR), ne nécessitent pas une grande puissance de calcul. Prenons un système linéaire :

$$Y = XA + E$$

Avec :

- A la matrice de prédiction et E la matrice de l'erreur.
- X le vecteur de variables d'entrée (X-variables) et Y le vecteur de sortie (Y-variables).

Résoudre le système consistera à calculer la matrice A puisque, lors de la phase de modélisation, les matrices X et Y sont connues. Dans le cas du MLR, il existe une solution simple aux sens des moindres carrés:

$$A = (X^T X)^{-1} X^T Y$$

On dit qu'on a fait une régression linéaire de Y sur X. Par contre on remarque que cette méthode ne permet pas le réglage de la complexité, ce qui est un inconvénient majeur de ce type de modélisation, notamment si la quantité des échantillons est faible. Par ailleurs, les phénomènes sont supposés être linéaires ce qui n'est pas le cas pour les capteurs de gaz.

4.3.2. Méthodes linéaires complexes, dites « factorielles »

Ces méthodes sont plus puissantes et capables de résoudre des systèmes complexes. Notons que les systèmes non linéaires peuvent souvent être traités par ces méthodes après linéarisation.

Ces méthodes sont toujours basées sur l'algèbre linéaire, elles sont puissantes et capables de résoudre certains problèmes non linéaires. Elles se basent sur l'analyse de facteurs, autrement dite analyse en composantes principales (PCA). Celle-ci consiste à transformer linéairement des variables liées entre elles (*variables X* dites "corrélées" à n dimensions) en nouvelles variables Z indépendantes les unes des autres en en diminuant le

nombre. Les nouvelles variables Z obtenues à h dimensions sont donc calculées par une combinaison linéaire des anciennes variables X , en cherchant à réduire le nombre de ces variables pour concentrer les informations utiles et en effectuant cette compression des données sans perdre d'informations (Smith, 2002). Les nouvelles variables Z sont appelées les « facteurs » ou « composantes principales » (voir figure 1-22).

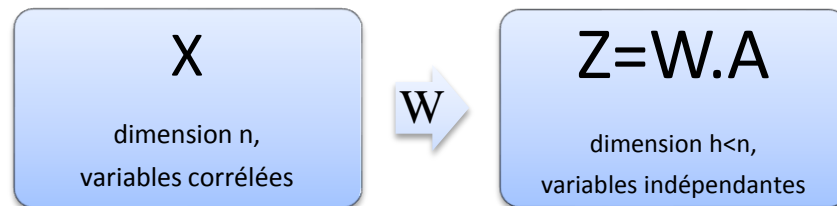


Figure 1-22 : Schéma d'un changement de base selon les méthodes de PCA.

On désigne par W la transformation de changement de base (ou le filtre matriciel à appliquer à X). L'utilisation des PCA dans les traitements de données aide à avoir une visualisation sur la corrélation entre les variables et la séparation des groupes de classe s'il en existe.

Il existe plusieurs méthodes pour déterminer cette matrice, d'où la différence entre les différentes méthodes de calcul factoriel. L'intérêt de cette transformation ou « compression » de données, est de grouper les informations utiles et diminuer le plus possible la présence du bruit. En appliquant cette méthode, les informations utiles dans les données seront groupées dans les premières valeurs, et le reste ne contient que le bruit inutile comme montré dans la figure 1-23. En appliquant des filtres matriciels, seules les valeurs intéressantes sont conservées. La complexité du modèle augmente avec le nombre de facteurs utilisés et est donc réglable.

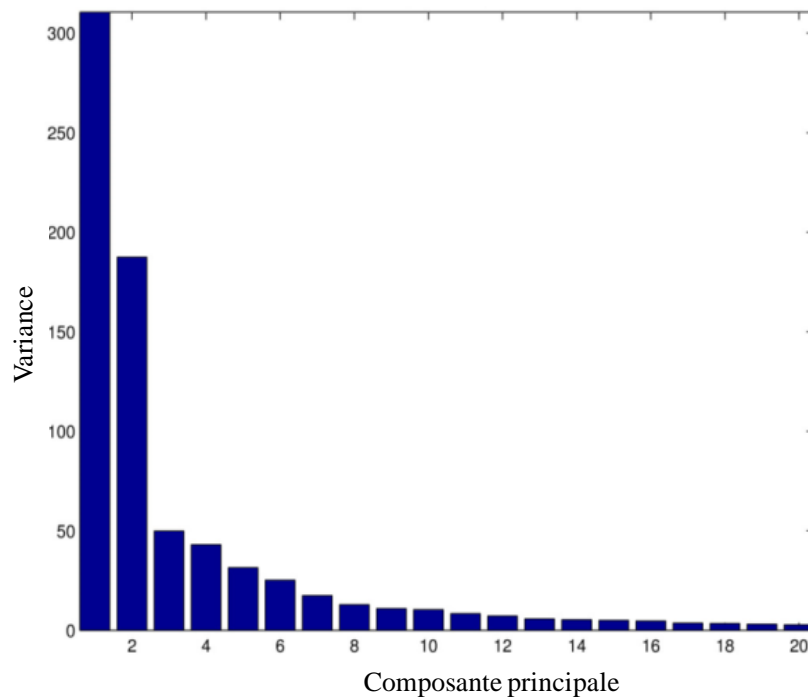


Figure 1-23 : variances des nouvelles variables après application d'une méthode PCA, on remarque que les informations nécessaires seront condensées dans les premières valeurs.

Il existe principalement deux méthodes d'analyse de facteur :

- **PCR (*Principal Component Regression*)** qui se base sur l'analyse en composantes principales (PCA) et donc sur l'étude des valeurs propres de la matrice XX' appelé la matrice de covariance, et donc sur les composantes principales des données X .
- **PLS (*Partial Least Squares*)** qui applique une méthode itérative pour trouver la matrice de changement de base en s'appuyant sur les corrélations entre les X et les Y variables.

Ces méthodes sont peu utilisées dans le cas des nez électroniques à cause de la non linéarités des capteurs utilisés, ce qui rend difficile et complexe de traiter les réponses non linéaire des capteurs avec des méthodes linéaires. Par contre, l'analyse en composantes principales peut être utilisée comme méthode de prétraitement de données dans les cas complexes afin de simplifier les modélisations.

4.3.3. Les réseaux de neurones artificiels

Les méthodes citées jusqu'à présent sont peu adaptées à la résolution des systèmes non linéaires, comme l'est la réponse d'un réseau de capteurs chimiques. Dans ce cas, on préférera utiliser les réseaux de neurones (NN : Neural Network) qui sont des algorithmes dits « de modélisation universelle ». Ce type d'algorithmes possède des termes et des définitions différentes de ceux rencontrés avant car ils ont pour origine la reconnaissance de forme et non la chimiométrie.

4.3.3.1. Structure générale du réseau

Un réseau de neurones artificiels est une méthode mathématique d'optimisation **itérative** dont le but est de prédire une solution d'un modèle physique. Il est constitué de nœuds de traitement des informations appelés neurones, qui sont connectés ensemble selon une architecture spécifique de l'algorithme appliqué. Chaque groupe de neurones forme une couche, et chaque neurone dépend de tous les neurones de la couche précédente comme le montre la figure 1-24. Cette figure montre les différentes couches d'un modèle : la couche d'entrée contenant les variables X , la (ou les) couche cachée contenant les variables Z , et la couche de sortie contenant les variables Y . Le nombre des couches cachées dépend de la complexité du modèle prévue, le système est ainsi appelé MLP (*MultiLayer Perceptron*) (Kamionka, 2005).

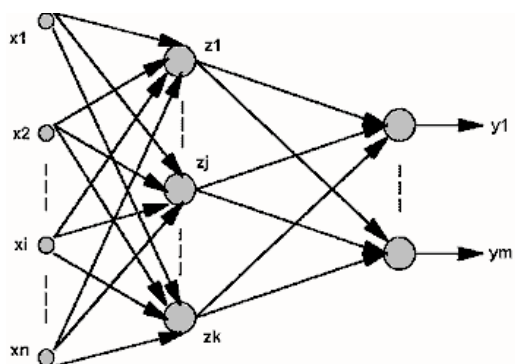


Figure 1-24 : architecture d'un réseau de neurones

La propagation des valeurs dans un réseau des neurones se fait de gauche à droite, chaque neurone est une combinaison linéaire de toutes les neurones de la couche précédente, les coefficients de cette combinaison sont appelés *poids synaptiques*. Dans chaque neurone, le résultat de la combinaison linéaire est transformé par une fonction non linéaire « dite fonction de transfert » comme indiqué sur la figure 1-25.

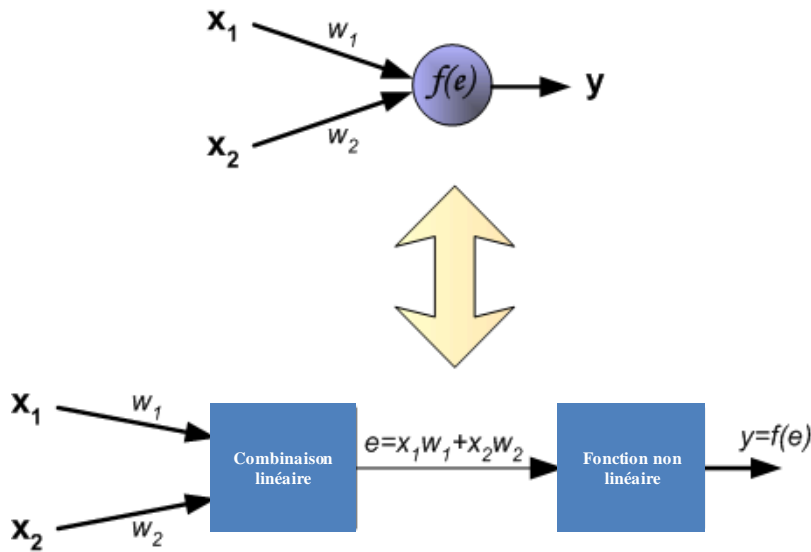


Figure 1-25 : Schéma d'un nœud de calcul (Kamionka, 2005).

Nous pouvons ainsi définir trois catégories de couches de neurones : couche d'entrées, couches intermédiaires ou cachées et couche de sortie. Prenons l'exemple d'un réseau MLP avec une seule couche cachée, nous pouvons ainsi calculer les valeurs de sortie du réseau comme suit :

$$Y_i = \beta_0 + \sum_i^m \beta_i f_i \left(\alpha_{i,0} + \sum_j^n \alpha_{i,j} x_j \right)$$

Avec :

- Y_i : la valeur de la sortie du neurone i de la couche de sortie de réseau.
- X_i : la valeur numéro i du vecteur des variables d'entrée.
- m : nombre des neurones dans la couche cachée
- n : nombre de variables d'entrée

- β_0 : l'offset à la sortie
- β_i : les coefficients de combinaisons à la sortie
- f_i : la fonction du transfert du neurone i de la couche cachée
- α_i : les coefficients de la combinaison à l'entrée de la couche cachée
- α_0 : offset du calcul

La fonction non linéaire utilisée est généralement une fonction bornée (*threshold function*) contenant un seuil d'activation k (Principe, et al., 2000):

$$f_t(x) = \begin{cases} 1 & \text{pour } x \geq k \\ 0 & \text{pour } x < k \end{cases}$$

Plusieurs types de fonctions non linéaires peuvent être utilisés comme montré dans la figure 1-26, la fonction la plus utilisée étant la tangente hyperbolique :

$$f_i = \tanh(x_i)$$

$$f_i = \frac{e^{x_i} - e^{-x_i}}{e^{x_i} + e^{-x_i}}$$

Nom	Valeur	Représentation
Seuil	$f(x) = 0$ si $x \leq 0$ $f(x) = 1$ si $x > 0$	
Seuil symétrique	$f(x) = -1$ si $x < 0$ $f(0) = 0$ $f(x) = 1$ si $x > 0$	
Linéaire saturée	$f(x) = 0$ si $x < 0$ $f(x) = x$ si $0 \leq x \leq 1$ $f(x) = 1$ si $x > 1$	
Linéaire	$f(x) = x$	
Sigmoïde	$f(x) = \frac{1}{1 + e^{-x}}$	
Tangente hyperbolique	$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	

Figure 1-26 : Fonctions non-linéaires utilisées dans les réseaux de neurones (Kamionka, 2005).

Si h_i dépasse le seuil d'activation k de la fonction on dit que le nœud est activé, sinon le nœud sera non activé et sa valeur devient nulle. Cette 2^{ème} étape peut être répétée plusieurs fois selon l'architecture du réseau.

Les algorithmes des réseaux de neurones sont caractérisés par 3 propriétés : leur architecture, leur fonction d'activation et la méthode d'apprentissage utilisée.

4.3.3.2. *Apprentissage*

Un des grands intérêts des réseaux de neurones, est leur aptitude à modéliser un système spécifique en passant par une phase d'apprentissage. C'est la partie équivalente à l'étalonnage dans les algorithmes multi-variables conventionnels. Le but de l'algorithme d'apprentissage est de déterminer par approximations successives les poids synaptiques (α_i et β_i) et les offsets (α_0 et β_0) d'un réseau de neurones afin que les valeurs de sorties et les valeurs cibles soient les plus proches possibles (au sens des moindres carrés). C'est un problème itératif, où chaque itération nous rapproche d'une solution optimale. Différents algorithmes d'optimisations existent, les plus utilisés sont : la rétro-propagation d'erreurs (*back-propagation*) ou l'algorithme de *Levenberg Marquard*. Ces algorithmes sont choisis soit par l'utilisateur soit automatiquement par les programmes de conception de réseaux de neurones.

Les réseaux de neurones peuvent être utilisés en apprentissage supervisé ou non supervisé (Otto, 2007). Avec l'apprentissage non-supervisé, seuls les vecteurs d'entrées sont utilisés pour classer des événements inconnus. Des algorithmes spécifiques pour de tels apprentissages existent comme les réseaux *Kohonen*. Ces réseaux sont connus pour leur nœuds qui s'auto-organisent, ils sont plutôt utilisés pour la classification des événements inconnus (Otto, 2007). Pour l'apprentissage supervisé, le réseau reçoit les informations additionnelles des variables Y « cibles » à prédire qui vont aider à définir les poids synaptiques du réseau. Dans le cadre de cette thèse, nous utiliserons les apprentissages

supervisés car nous cherchons à déterminer des valeurs qualitatives ou semi-quantitatives en relation avec la présence de polluants dans l'atmosphère.

L'objectif de l'apprentissage est de minimiser les différences entre les valeurs de sorties du système et les valeurs désirées. Donc, pour quantifier le degré d'apprentissage, on utilise une fonction d'erreur. La fonction la plus utilisée est le **RMSEP** (*Root mean square error of prediction*) **c'est la racine carrée de la moyenne des carrés des erreurs de prédiction** (26). La fonction est écrite de la façon suivante :

$$RMSEP = \sqrt{\frac{\sum(Y - Y')^2}{n}}$$

Avec :

- Y et Y' sont respectivement les valeurs attendues et prédites par le modèle.
- n est le nombre total d'observations calculées

Au cours de l'apprentissage, le RMSEP des échantillons d'étalonnage va décroître au fur et à mesure que la modélisation devient plus complexe et plus précise. Pour éviter le problème de sur-modélisation, le réglage de la complexité du modèle se fait en limitant le nombre d'itérations. Il est pour cela nécessaire de tester les valeurs de modélisation durant les itérations d'apprentissage avec un échantillon de données distinct de celui utilisé pour l'apprentissage. On parle d'échantillons de « validation ». Le modèle construit à chaque itération d'apprentissage sera testé avec les échantillons de validation, les résultats sur ces échantillons vont ainsi intervenir quelque peu pour construire le modèle. Après avoir construit le modèle, un autre échantillon de données n'appartenant ni à l'apprentissage ni à la validation devra donc être utilisé pour qualifier les performances finales du modèle. On parle d'échantillon de « test ». Au cours de la modélisation, il est indispensable que les échantillons d'apprentissage, de validation et de test soient indépendants les uns des autres. Il est important

aussi d'avoir une indépendance temporelle entre les données d'apprentissage, validation et test.

4.4. Système multi-capteurs et analyse multi-variables

Les systèmes multi-capteurs couplent des capteurs chimiques partiellement sélectifs (spécialement des capteurs à semi-conducteurs) avec un algorithme de reconnaissance de forme à base d'analyses multi-variables pour la recherche d'« empreintes chimiques » (Holmberg, et al., 1995) (Shaffer, et al., 1999). Le choix de l'algorithme est un élément important de ces systèmes, il s'agit donc pour les chercheurs de développer des méthodes de reconnaissance de forme en adéquation avec l'application visée (Fu, et al., 2007).

4.4.1. Système multi-capteurs et réseaux de neurones

Les réseaux de neurones sont largement utilisés dans les algorithmes de reconnaissance de forme. Que l'application soit simple, comme la détection des concentrations des gaz dans des mélanges binaires (Martin, et al., 2001), ou plus complexe, comme la détection d'altération d'huile (Hai, et al., 2006) ou encore la discrimination entre différentes espèces aromatisantes (Brezmes, et al., 1997). L'utilisation des réseaux de neurones offre des avantages majeurs en termes de capacité de prédiction, de traitement des non linéarités, de robustesse, et de limitation des phénomènes de sur-modélisation (Szecowka, et al., 2011). Ils sont largement utilisés par les chercheurs dans les domaines des nez électroniques et des réseaux de capteurs (Sommer, et al., 1995) (Gardner, et al., 1999).

4.4.2. Prétraitement et normalisation des données

Un modèle idéal combine une précision, une vitesse de calcul, une facilité d'apprentissage et une robustesse face à des interférents au sens mathématique (Natale, et al., 1995). En réalité, il n'y a aucun algorithme de reconnaissance de formes capable d'assurer l'ensemble des propriétés citées précédemment. Certains travaux ont démontré l'importance

d'un prétraitement des données qui précède l'algorithme lorsque le nombre de capteurs et d'espèces gazeuses à identifier est grand (Hines, et al., 1999), et que les signaux de sortie des capteurs sont difficiles à traduire en données utiles. Cette partie de transition n'étant pas un vrai traitement de données, elle est considérée parfois comme une partie de l'algorithme de reconnaissance de forme. Ce prétraitement peut varier d'une simple normalisation des signaux des capteurs et filtrage des erreurs (Martin, et al., 2001), jusqu'à l'utilisation de méthodes linéaires plus complexes comme les ACP (analyse en composantes principales) pour faire un pré-classement des données (Lozano, et al., 2008) avant d'alimenter le réseau de neurones. L'intérêt de ce prétraitement est d'éliminer considérablement les erreurs de mesures et les informations inutiles et, dans les cas extrêmes, de compresser les données en diminuant le nombre des variables à modéliser (Negri, et al., 2001) (Kuske, et al., 2005).

Voici un exemple de prétraitement de données, tiré des travaux de Miguel Martin, sur la discrimination de 2 espèces gazeuses (NO_2 et CO) (Martin, et al., 2001). Les valeurs d'entrées ont été normalisées suivant l'équation :

$$x_i = \log \left(R_{\text{gaz}} / R_{\text{ref}} \right)$$

Avec :

- x_i est la valeur attribuée au capteur i .
- R_{gaz} et R_{ref} : les résistances du capteur i respectivement pour le gaz spécifié et un air de référence.

Il existe d'autres techniques de normalisation utilisées comme prétraitement de données (Romain, et al., 2000). Ainsi on trouve la méthode utilisée par AC Romain et al (Romain, et al., 2005), où la résistance d'un capteur sera divisée par la somme quadratique des résistances de l'ensemble des capteurs du système :

$$x_i = \frac{R_i}{\sqrt{\sum R_j^2}}$$

Avec :

- x_i les valeurs attribuées au capteur i ;
- R_i la résistance du capteur i ;
- R_j les résistances des capteurs du système incluant le capteur R_i .

L'intérêt de cette méthode est de normaliser les résistances (ou conductivités) des différents capteurs d'un système en éliminant les informations communes et en amplifiant les différences entre eux. Cette méthode est utilisée comme un prétraitement pour alimenter un algorithme qui cherche à identifier les différences entre les signaux des capteurs et non pas les corrélations entre eux.

4.4.3. Systèmes multi-capteurs et qualité de l'air.

Plusieurs études ont été menées avec des systèmes multi-capteurs appliqués à la qualité de l'air (Tsujita, et al., 2005) (Romain, et al., 2005) (Lee, et al., 2001). En général, ces études peuvent être divisées en deux catégories. Tout d'abord, on trouve celles effectuées en laboratoire. Elles concernent le suivi d'une ou plusieurs espèces gazeuses toxiques dans une atmosphère synthétique en laboratoire, comme la mesure de concentrations en H_2S et NO_2 en mélange gazeux (Helli, et al., 2004) (Yang, et al., 2007) ou la détection de mélanges de COV par des capteurs à semi-conducteurs (Szcurek, et al., 2010) (Schweizer-Berberich, et al., 1998) (Saeed, et al., 2009). Par ailleurs, d'autres études sont faites directement sur le terrain, en sites pollués, où un ou plusieurs capteurs sont placés dans des sites spécifiques pour détecter la présence d'une ou plusieurs espèces gazeuses dans l'atmosphère. Nous classons ces études dans le tableau 1-10 :

Tableau 1-10 : Liste des travaux effectués sur la détection de polluants avec des réseaux de capteurs.

Catégorie	Ref	Application visée	Méthodes d'analyses utilisées
Etude en laboratoire	(Lee, et al., 2001)	Discrimination des COVs (10-1000 ppm)	PCA-ANN
	(De Vito, et al., 2009)	Mélange CO-NO _x	NN-Automatic bayesian regulation
	(Negri, et al., 2001)	Mélange CH ₄ -CO-Isobutane-ethanol(concentration totale: 500-2500 ppm)	ANN
	(Aleixandre, et al., 2004)	Mélange NO ₂ (0,5-2 ppm), CO(50-150 ppm), Toluene (50-150 ppm), Octane (50-150 ppm)	PCA, ANN, PNN, MLP, GZ
	(Chaiboun, et al., 2007)	Mélange NH ₃ , CO, CH ₄ , C ₃ H ₈ , Ethanol (1-1000 chacun)	Logarithmic model
	(Yang, et al., 2007)	Mélange H ₂ S NO ₂ (0-9 ppm)	ANN-MLP
	(Martin, et al., 2001)	Mélange CO (0-45 ppm) NO ₂ (0-125 ppb)	ANN
	(Helli, et al., 2004)	Mélange H ₂ S (1-11 ppm) NO ₂ (1-5 ppm) avec humidité variable	LDA
	(Szczurek, et al., 2010)	Reconnaissance Benzene Xylene Toluene	PCA-LDA
	(Strobel, et al., 2004)	Mélange H ₂ S, NO ₂ , SO ₂	LDA
	(Szczurek, et al., 2011)	Mélange des différents COVs	LDA
	(Getino, et al., 1999)	Mélange des différents polluants (NO _x , Benzène, SO ₂)	ANN
	(Zampolli, et al., 2004)	Mélange CO et NO ₂	Fuzzy ANN
	(Saeed, et al., 2009)	Mélange des différents COVs	ANN
Etude sur terrain	(Tsujita, et al., 2005)	Suivi NO ₂ dans l'atmosphère	Auto-calibration
	(Kamionka, et al., 2006)	Suivi de NO ₂ Et O ₃ dans l'atmosphère	ANN

D'autres études ont été menées sur la stabilité des signaux (Chaiboun, et al., 2007) et sur la dérive des réponses des capteurs, en cherchant des solutions analytiques à ces problèmes, au risque de complexifier les systèmes conçus (Brasan, et al., 1995) (Padilla, et al., 2010) (Flitti, et al., 2008). Par exemple, W. Tsujita et al. ont proposé une méthode d'auto-calibration pour résoudre le problème de dérive des capteurs durant leurs travaux sur le suivi de la concentration de NO₂ dans l'atmosphère (Tsujita, et al., 2005). Ainsi leur système a été

installé dans des stations de mesure de la qualité de l'air et leur méthode d'auto-calibration est basée sur la comparaison des niveaux de NO₂ mesurés par les analyseurs des stations de surveillance de l'air et les concentrations calculées à partir des signaux des capteurs. Cette auto-calibration est parvenu à diminuer les erreurs causées par la dérive des capteurs, par contre, elle nécessite un accès aux informations sur les concentrations de NO₂ dans les sites surveillés. Par ailleurs, A.C. Romain et J. Nicolas ont étudié la dérive des capteurs de type TGS pendant 7 ans. Ils ont montré l'importance de développer un système capable de compenser cette dérive d'une part et de tenir compte de la non reproductibilité de fabrication des différents capteurs de même type (Romain, et al., 2010). Même si l'on remplace un capteur par un autre, son comportement, et donc sa dérive, ne sera jamais strictement identique à celui d'origine du même type et provenant du même fournisseur. Il est donc nécessaire de recalibrer le système à chaque échange de capteur (Tomic, et al., 2004). Dans leurs travaux, ils ont réussi à développer un modèle correcteur qui estime la pente de la dérive d'un capteur, et calcule un facteur multiplicatif pour chaque capteur. Même avec une telle correction, un calibrage trimestriel sous un gaz étalon est nécessaire pour maintenir l'efficacité du modèle.

D'autre part, des études ont montré que la discrimination multi-source est possible avec les réseaux de capteurs (Nicolas, et al., 2006). C'est par exemple le cas dans les problématiques de nuisances olfactives : lorsqu'un site récepteur subit l'influence de plusieurs sources, l'identification de la source majoritaire à l'origine de la nuisance permettra d'appliquer les actions correctives ad hoc. C'est l'une des applications des réseaux de capteurs les plus largement usitées. On peut par exemple identifier le type de polymères utilisé dans la construction des voitures (Morvan, et al., 2000), ou encore détecter la fumée de cigarettes dans un objectif de sécurité (Ködderitzsch, et al., 2005). Pour ce qui concerne la localisation géographique d'une ou plusieurs sources à l'origine d'une nuisance, l'utilisation d'un réseau

de capteurs associé à un robot mobile permettant de se diriger vers la source a été étudiée (Trincavelli, et al., 2008). Cette approche présente beaucoup de difficultés en raison de la nature géographique de l'environnement (présence éventuelle d'obstacles), et de la prise en compte de sources ponctuelles imprévues. Un maillage du terrain par des réseaux de capteurs stationnaires permet de s'affranchir des aléas liés à la géographie et pallier aux sources instantanées intempestives, mais ce système de réseaux fixes doit prendre en considération les variables météorologiques, telles que la vitesse du vent (Cai, et al., 2007).

Toutes ces études montrent la polyvalence d'un réseau de capteurs, un seul réseau de capteurs peut avoir diverses applications simplement en variant la cible de l'analyse multivariable utilisée.

5. Conclusion et objectifs de la thèse

L'analyse de la pollution atmosphérique utilise des réseaux de stations de mesures de la qualité de l'air mettant en œuvre des analyseurs basés sur des techniques d'échantillonnage qui nécessitent souvent des équipements complexes, lourds, et coûteux. Or, pour les applications de détection de pollution où la sélectivité et la précision ne sont pas une priorité, l'utilisation de ces techniques peut s'avérer surdimensionnée. Par ailleurs, l'utilisation des capteurs électroniques de gaz peut fournir une réponse « *low cost* » à la problématique de surveillance de la qualité de l'air. Nous cherchons ainsi à développer un système à base de capteurs de gaz pour le suivi de la qualité de l'air en assurant un flux d'informations fiable et robuste sur leur environnement gazeux sans pour autant être tenu de fournir la composition physico-chimique précise de l'air ambiant.

Ainsi, l'objectif de ce travail est le développement d'un système permettant l'identification de différents types de pollutions atmosphériques ainsi que l'évaluation du niveau de cette pollution et la validation du système développé sur un site multi-influencé. Il

s'agira donc dans un premier temps de réaliser un réseau de capteurs de gaz (en privilégiant les capteurs à semi-conducteurs), dans le but de développer un modèle mathématique (Réseaux de neurones, ACP...) permettant de traiter les données issues des signaux des capteurs afin d'obtenir une qualification de la pollution atmosphérique et une quantification de son niveau. Ce sujet se distingue de la plupart des études menées jusqu'à présent puisque la validation s'effectuera, non pas en laboratoire sur des mélanges de gaz synthétiques et relativement simples et dans des conditions bien contrôlées (Tsujita, et al., 2005) (Negri, et al., 2001) (Martin, et al., 2001) (Szczurek, et al., 2010), mais sur le terrain, avec des mélanges de gaz complexes en conditions réelles, et en testant différentes catégories de sites.

Dans cette thèse nous avons opté pour l'utilisation de capteurs à semi-conducteurs commerciaux pour leur facilité d'intégration dans un système portable, leurs faible coût et consommation électrique, leur bonne sensibilité et enfin pour profiter des derniers progrès technologiques dans ce domaine. Afin d'améliorer les performances du système, ce travail s'effectuera donc en 2 phases, décrites ci-après.

La première partie du travail consistera donc à tester puis choisir judicieusement les capteurs les plus adaptés à l'application terrain dans le domaine de la pollution atmosphérique, à les mettre en œuvre dans un système autonome portable, à développer un banc de test de validation au laboratoire en utilisant des gaz synthétiques afin d'établir une procédure de vérification de bon fonctionnement des cellules réalisées. Plusieurs systèmes multi-capteurs seront ainsi testés simultanément pour assurer une répétabilité et reproductibilité des réponses des capteurs, ainsi que des modèles mathématiques associés.

Cette partie de ce travail se déroule dans le département MICC au centre SPIN de l'EMSE en raison de l'expérience acquise au laboratoire dans le domaine des capteurs chimiques de gaz, spécialement en ce qui concerne les capteurs à semi-conducteurs : les travaux sur les capteurs

à semi-conducteurs pour la détection des gaz ont débuté dans le département MICC du centre SPIN dès 1981 avec C. Pijolat (Pijolat, 1986). Par la suite, l'utilisation de ces capteurs sur le sujet de la détection multi-gaz a commencé avec N. Perdreau en 2000 (Perdreau, 2000) qui a étudié la capacité d'un système constitué de plusieurs de ces capteurs et a effectué la détection quantitative de gaz en utilisant les techniques de réseaux de neurones. Ces travaux ont été suivis par ceux de M. Kamionka (Kamionka, 2005) en ciblant le sujet de la qualité de l'air ambiant et en utilisant un système multi-capteurs couplé avec les réseaux de capteurs pour le suivi continu de l'ozone dans une pollution atmosphérique. Une partie de ces travaux a été dédiée au développement des capteurs à semi-conducteurs. Le centre SPIN est donc particulièrement compétent dans la compréhension des différents phénomènes physiques régissant le fonctionnement de ces capteurs.

La seconde partie de ce travail sera menée au département Chimie et Environnement de l'Ecole des Mines de Douai compte tenu de son expérience dans le domaine de la mesure et de la surveillance de la qualité de l'air. Elle se déroulera en deux phases :

- Mesures sur site avec les systèmes multi-capteurs conçus et validés à Saint Etienne, placés sur différents sites de mesures. Pour cela, des contacts ont été pris avec les AASQA qui sont en charge de la surveillance de la qualité d'air en France et en particulier les AASQA qui sont en charge de cette surveillance en Picardie et en Région Nord/Pas-de-Calais afin de pouvoir disposer des données annuelles de concentration des polluants. Ces données seront statistiquement exploitées pour définir et identifier les différents types de pollution afin de les comparer avec les réponses des capteurs. Développement d'un modèle de comportement basé sur les méthodes d'analyses multi-variables (PLS, NN...) qui sera construit à partir des mesures délivrées par les réseaux de capteurs pour converger vers des signatures des différents types de pollutions

établies à partir des données de concentrations de polluants disponibles dans les stations de mesure.

Finalement la validation de ce dispositif intégrant le modèle de comportement sera effectué dans un site différent de ceux intégrés dans la phase d'apprentissage : le système devra identifier le type de la (ou des) pollutions existantes sur un site « inconnu ». Cette partie du travail nécessitera l'optimisation du système réalisé soit par l'ajout de certains capteurs soit par l'amélioration du modèle établi.

Chapitre II :

Développement, mise en œuvre et validation d'un réseau de capteurs.

1. Introduction

L'objectif de cette thèse est la conception d'un système multi-capteurs portable et autonome pour la détection des différents types de pollution atmosphérique.

Dans ce chapitre est présentée la méthodologie adaptée pour concevoir le système de détection des gaz, en partant du choix des capteurs, jusqu'à la gestion de sa portabilité et de son autonomie. Seront également détaillés, les bancs de caractérisation des performances métrologiques des réseaux de capteurs développés durant ce travail de thèse, les différentes performances des capteurs obtenues, et le mode opératoire des études de calibrage et de performance des capteurs.

2. Choix des capteurs

Nous avons décidé de travailler avec des capteurs à semi-conducteurs commerciaux en raison de leur simplicité de mise en œuvre et de leur coût faible. De plus, ce choix a été guidé par le contexte de suivi de la pollution atmosphérique présenté dans le chapitre 1, car les limites de détection sont a priori acceptables au regard des niveaux attendus, et la combinaison de plusieurs capteurs associés à une modélisation ad hoc devrait permettre d'établir une signature spécifique d'une pollution. Le marché des fabricants des capteurs de gaz propose plusieurs types de produits basés sur différentes technologies dont quelques exemples typiques sont présentés dans le tableau 2-1. L'étude bibliographique a montré que les capteurs les plus utilisés sont les capteurs TGS construits par Figaro, leader historique mondial du marché.

Tableau 2-1 : Liste des constructeurs de capteurs de gaz. En gras capteurs de type semi-conducteurs.

Entreprise	Type de capteur	Produit	Familles des composés détectés
Alphasense	Electrochimique	α sense	CO, H ₂ S, NO ₂ , NO, SO ₂
	Pid	Pid	VOC
E2v	Semi-conducteur	MICS	VOC, NH₃, CO, NO₂, O₃
	Electrochimique	EC	NO ₂ , NO, H ₂ S, CO, SO ₂
	Infrarouge	IR	CO ₂
	Conductivité thermique (Pellistor)	VQ	CH ₄ , H ₂ S
Figaro	Semi-conducteur	TGS	CH₄, VOC
	Electrochimique	CDM	O ₂ , CO ₂ , H ₂
Fis	Semi-conducteur	FIS	VOC
Membrapor	Electrochimique	MEMBRAPOR	O ₂ , SO ₂ , O ₃ , NH ₃ , NO, NO ₂ , VOC
Microsense	Semi-conducteur	MSGS	O₂, H₂, CO, NO_x, CH₄
	Catalytique	MCGS	H ₂ , CH ₄ , VOC
	Conductivité thermique	MTCS	CH ₄ , CO ₂

Notre choix s'est porté sur deux fabricants pour des raisons de notoriété et de disponibilité de leurs produits: Figaro (Japon) et MICS (Suisse), chacun proposant toute une gamme couvrant nos objectifs d'applications.

2.1. Propriétés des capteurs

L'ensemble des capteurs choisis doit permettre de couvrir toutes les espèces ou familles d'espèces susceptibles d'être présentes dans une pollution atmosphérique. La nature des polluants réglementés (O_3 , NO_2 , NO , CO) est un des premiers critères à considérer, car leur présence dans l'atmosphère donne une indication directe sur le niveau ou sur le type de pollution atmosphérique, mais surtout leur mesure étant obligatoire, il sera aisé d'accéder à des bases de données informant sur les niveaux de concentrations mesurés et permettant d'établir une comparaison avec les données issues des capteurs. En second lieu, les COV sont des gaz importants à cibler parce qu'ils peuvent être également spécifiques d'un type de pollution présent dans l'air ambiant (voir chapitre 1 partie 1.1.). Enfin, nous avons complété notre choix par des capteurs capables de détecter des gaz non réglementés comme NH_3 , H_2S , et les alcanes car ils sont eux aussi spécifiques de certains types de pollutions plutôt d'origine industrielle mais aussi trafic. Le Tableau 2-2 répertorie les capteurs choisis avec leur gaz cibles et leurs gammes de détection données par leurs constructeurs. Les constructeurs donnent pour chaque capteur la résistance de la partie sensible et celle du chauffage ainsi que les tensions respectives associées (tension de chauffage V_H , et tension de mesure V_C). Le Tableau 2-2 récapitule les caractéristiques constructeur pour chaque capteur choisi initialement dans la conception du système multi-capteurs ainsi que la résistance R_L de mesure choisie.

Nous remarquons sur ce tableau que les gammes de concentrations indiquées par les constructeurs sont le plus souvent bien au-delà des concentrations rencontrées en pollution atmosphérique, nous devons donc vérifier par des tests en laboratoire la possibilité de détecter des faibles concentrations.

Tableau 2-2 : Choix des capteurs.

Capteur	Gaz cibles	Gamme de détection (constructeur)	V _C	V _H	R _L	Consommation électrique
MICS5525	Bonne sensibilité pour CO, filtre charbon actif pour éliminer l'éthanol et COV.	1-1000 ppm	5 V	2.5 V	1 MΩ	76 mW
MICS2710	NO ₂	0,5-5 ppm	2.5 V	1.7 V	10 KΩ	43 mW
MICS2611	O ₃	0,1-1 ppm	5 V	2.5 V	4.7 KΩ	80 mW
TGS2600	Large spectre des contaminants de l'air	500-10000 ppm	5 V	5 V	10 KΩ	280 mW
TGS2602	COV-Odeurs-H ₂ S	30-10000 ppm	5 V	5 V	10 KΩ	280 mW
TGS2611	Gaz combustibles - Méthane	500 – 10000 ppm	5 V	5 V	10 KΩ	280 mW
TGS2620	Alcool et vapeurs de solvant	50 – 10000 ppm	5 V	5 V	10 KΩ	280 mW
MICS5521	CO-COV et hydrocarbures	0,1 – 100 ppm	5 V	2.5 V	47 KΩ	76 mW
MICS5914	NH ₃	300-40000 ppm	5 V	2.5 V	10 KΩ	66 mW

2.2. Mise en œuvre des capteurs à semi-conducteurs

Ces capteurs semi-conducteurs nécessitent deux sources de tension continues, une pour la mesure de résistance de la couche sensible, et l'autre pour alimenter la résistance de chauffage. La tension de chauffage doit correspondre à la valeur préconisée par le constructeur afin que la couche sensible réponde au(x) gaz ciblé(s) en fonction d'une température donnée. Le montage électronique correspondant à ce type de capteur est simple, comme montré dans la figure 2-1.

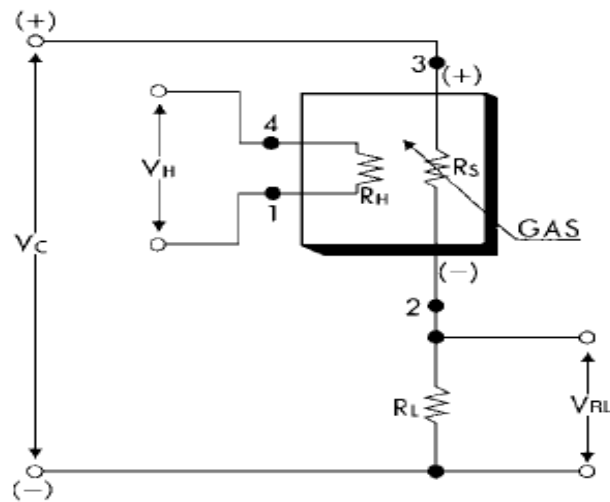


Figure 2-1 : Montage électronique pour la mise en œuvre d'un capteur semi-conducteur. Source : www.figaro.co.jp

Dans cette figure nous trouvons :

- R_H : résistance de chauffage
- R_S : résistance de la couche sensible, dépend de l'environnement gazeux
- R_L : résistance de mesure
- V_H : tension de chauffage
- V_C : tension de polarisation de la couche sensible
- V_{RL} : tension mesurée aux bornes de la résistance de charge R_L

La mesure de la conductance $G_S = \frac{1}{R_S}$ (1) de la couche sensible est ainsi réalisée grâce à un pont diviseur de tension incluant la résistance de mesure R_L :

$$R_S = \frac{V_C - V_{R_L}}{V_{R_L}} * R_L \quad (2)$$

Cette conductance G_S varie en fonction des concentrations des espèces gazeuses dans l'environnement du capteur.

2.3. Systèmes de mesure d'humidité et de température du gaz

Comme déjà expliqué (voir chapitre 1 partie 3.1.3.), les capteurs semi-conducteurs sont sensibles aux changements de la température et de l'humidité de leur environnement gazeux. Même si dans le cadre de ce travail de thèse nous sommes toujours parvenus à travailler dans des conditions de température et d'humidité relative contrôlées, soit en laboratoire, soit dans les stations de mesure climatisées des associations de mesure de la qualité de l'air, il nous a paru nécessaire de prévoir la mesure de ces deux grandeurs d'influence et de suivre leurs variations afin de corriger leurs effets si les réseaux de capteurs devaient être utilisés par la suite dans des conditions moins maîtrisées. C'est la raison pour laquelle deux capteurs de température et d'humidité relative (HR) ont été intégrés dans notre dispositif pour le compléter.

La mesure de la température se fait à l'aide d'un capteur de température de type LM35. Ce capteur comporte 3 contacts, une masse, une alimentation comprise entre 4 et 20V et une sortie comme montré dans la figure 2-2. La tension de sortie de ces capteurs est proportionnelle à la température et sa sensibilité est de 0,01V/°C. Un calibrage est nécessaire pour corriger l'offset de tension à une température de 0°C.

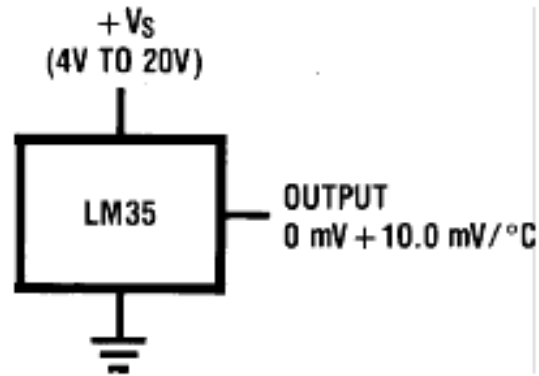


Figure 2-2 : Schéma du montage du capteur de température LM35.

La courbe d'étalonnage de ce capteur a été obtenue en chambre climatique pour des températures comprises entre 0 et 40°C, représentatives des futures conditions d'utilisation du réseau de capteurs. La figure 2-3 illustre cette courbe, une relation linéaire est claire entre la température du capteur et la température ambiante, et l'offset de tension à 0°C est ainsi ici de 5,3°C.

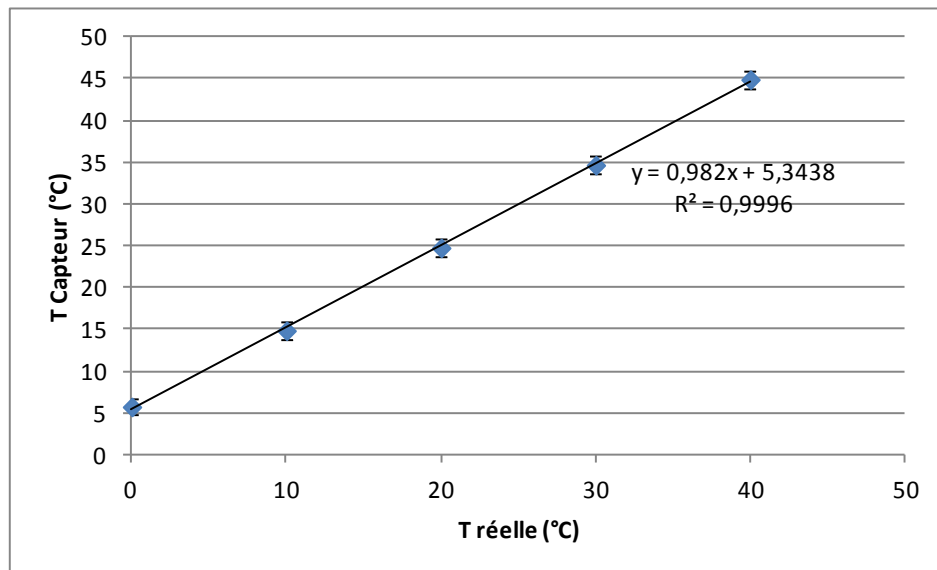


Figure 2-3 : Etalonnage du capteur de température

Pour la mesure de l'humidité de l'air, un capteur capacitif a été utilisé : le HS1101LF (*Measurement Specialties*). Le condensateur, élément sensible de ce capteur, est associé à un oscillateur basé sur un circuit intégré « timer » de type 555 comme montré dans la figure 2-4.

On mesure alors en sortie un signal carré de fréquence inversement proportionnelle à la capacité C_1 .

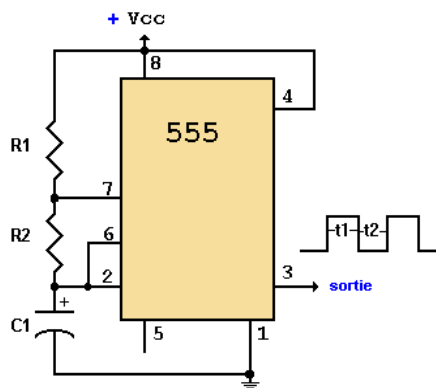


Figure 2-4 : Schéma de montage de l'oscillateur conditionnant le signal du capteur d'humidité C_1

La relation entre la fréquence et la valeur de la capacité est donnée par :

$$C_1 = \frac{1,44}{(R_1 + 2 * R_2) * F} \quad (3)$$

Si les résistances utilisées R_1 et R_2 sont identiques et égales à R , l'équation 3 devient :

$$C_1 = \frac{1.44}{3 * R * F} \quad (4)$$

Avec :

- R : la valeur de résistances utilisées, nous utilisons une résistance de 470 k Ω
- F : la fréquence du signal de sortie du circuit en Hz
- C_1 : la capacité du capteur d'humidité en pF.

D'après le constructeur, la relation entre la capacité de ce capteur et l'humidité relative est linéaire, il en est donc de même pour la relation entre humidité relative et fréquence de sortie de l'oscillateur. Par conséquent, un simple étalonnage permet d'accéder à l'équation linéaire pour le calcul de l'humidité directement à partir de la fréquence de sortie de l'oscillateur comme montre la figure 2-5.

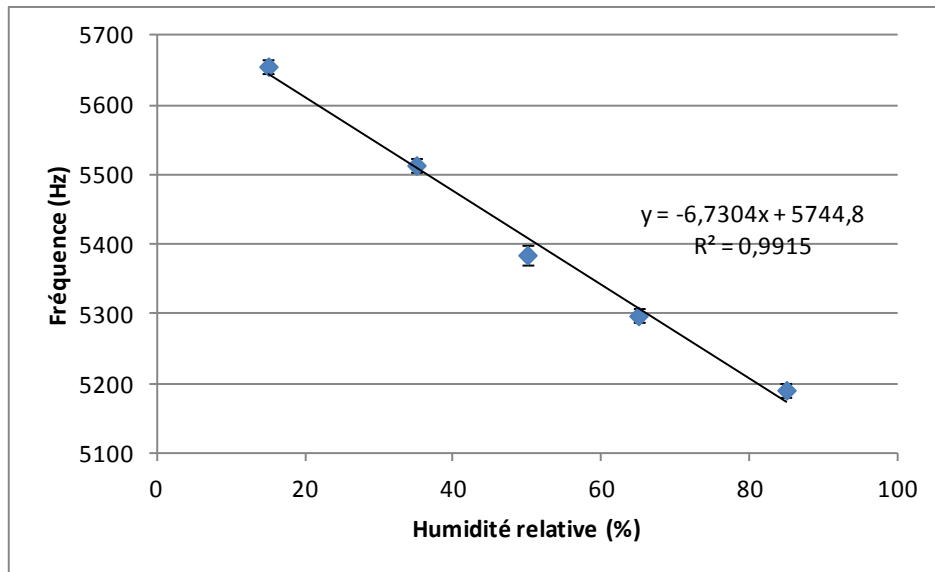


Figure 2-5 : Courbe d'étalonnage du capteur d'humidité relative.

Afin d'étudier l'influence de l'humidité sur la réponse des capteurs, il est nécessaire d'introduire la notion d'humidité absolue, car le couple [humidité relative/température] n'est pas toujours suffisant pour expliquer tous les phénomènes observés (voir paragraphe 4.2.). L'humidité absolue correspond à la concentration molaire en vapeur d'eau, tandis que l'humidité relative est le rapport entre l'humidité absolue dans l'air et l'humidité absolue saturante de l'air. L'humidité absolue saturante est la quantité maximale en vapeur d'eau que l'air peut contenir sans condensation. L'humidité absolue peut être calculée à partir de l'humidité relative en utilisant les équations suivantes :

$$H_a = \frac{H_{asat} * H_r}{100} \quad (5)$$

$$H_{asat} = \frac{P_{sat} * M_{eau}}{P_{atm} * M_{air}} \quad (6)$$

$$P_{sat} = \exp\left(13,7 - \frac{5120}{T}\right) \quad (7)$$

Avec :

- H_{asat} , H_a , H_r : l'humidité absolue saturante, l'humidité absolue et l'humidité relative respectivement.
- M_{eau} , M_{air} : la masse molaire de l'eau (18g/mole) et de l'air (28.9g/mole) respectivement.

- P_{sat} , P_{atm} : la pression de vapeur d'eau saturante approximée par la formule de Rankine (7), et la pression atmosphérique respectivement, les deux en bar (10^5 hPa).
- T : température ambiante en K.

Par conséquent, nous pouvons suivre l'humidité absolue de l'air à tout moment en utilisant la valeur de l'humidité relative (HR) obtenue à partir du capteur d'humidité, et de la valeur de la température de l'air obtenue à partir du capteur de la température.

2.4. Conception de la cellule de capteurs en laboratoire

Pour assurer le bon fonctionnement des capteurs, plusieurs prototypes de cartes électroniques ont été développés afin d'assembler, de supporter et d'assurer les connections électriques entre les différentes parties de notre système. Dans un premier temps, notre système a été intégré dans une cellule en polymère. Mais au cours des expériences nous nous sommes rendu compte des effets négatifs susceptibles de se produire avec une telle cellule, comme l'adsorption des gaz par les polymères ou inversement la désorption de gaz contenus dans le polymère. Pour cette raison, une 2nde cellule en verre cette fois, a été réalisée afin d'éliminer tout problème d'adsorption. Cette dernière est équipée d'une carte circuit imprimé qui peut supporter 9 capteurs et assure leur alimentation en tension de chauffage et de mesure. Elle contient aussi les systèmes de mesures de température et d'humidité relative. L'acquisition de tous les signaux est rendue possible par un connecteur de type DB25. Trois exemplaires identiques de ces cellules ont été réalisés. Finalement, ces 3 cellules seront utilisées en parallèle pour évaluer la reproductibilité des signatures entre les différentes cellules. Nous montrons dans l'annexe A page 193 des photos des systèmes conçus.

3. Développement de bancs de test

Afin d'évaluer et valider les performances des systèmes multi-capteurs ainsi réalisés dans des conditions environnementales les plus proches possibles des futures applications envisagées sur le terrain, des bancs de tests capables de tester en parallèle les 3 cellules multi-

variables ont été élaborés. Leur but est d'assurer le passage d'un débit d'air constant et contrôlé simultanément dans chacune des cellules.

3.1. Banc de test de validation

Pour valider à l'aide de mélanges gazeux le système multi-capteurs et d'acquérir les signaux électriques issus des différents capteurs, nous avons construit un banc de test automatisé et adapté à plusieurs types de gaz. Ce banc de test a été élaboré de sorte à envoyer une ou plusieurs espèces gazeuses présentes habituellement dans l'air ambiant lors des épisodes de pollution vers nos cellules. Les gaz choisis sont : CO, CH₄, toluène, éthanol, NO₂, H₂S, O₃. Ce flux de gaz est partagé et traverse les cellules multi-capteurs contenant les réseaux de capteurs. La figure 2-6 illustre le schéma de ce banc.

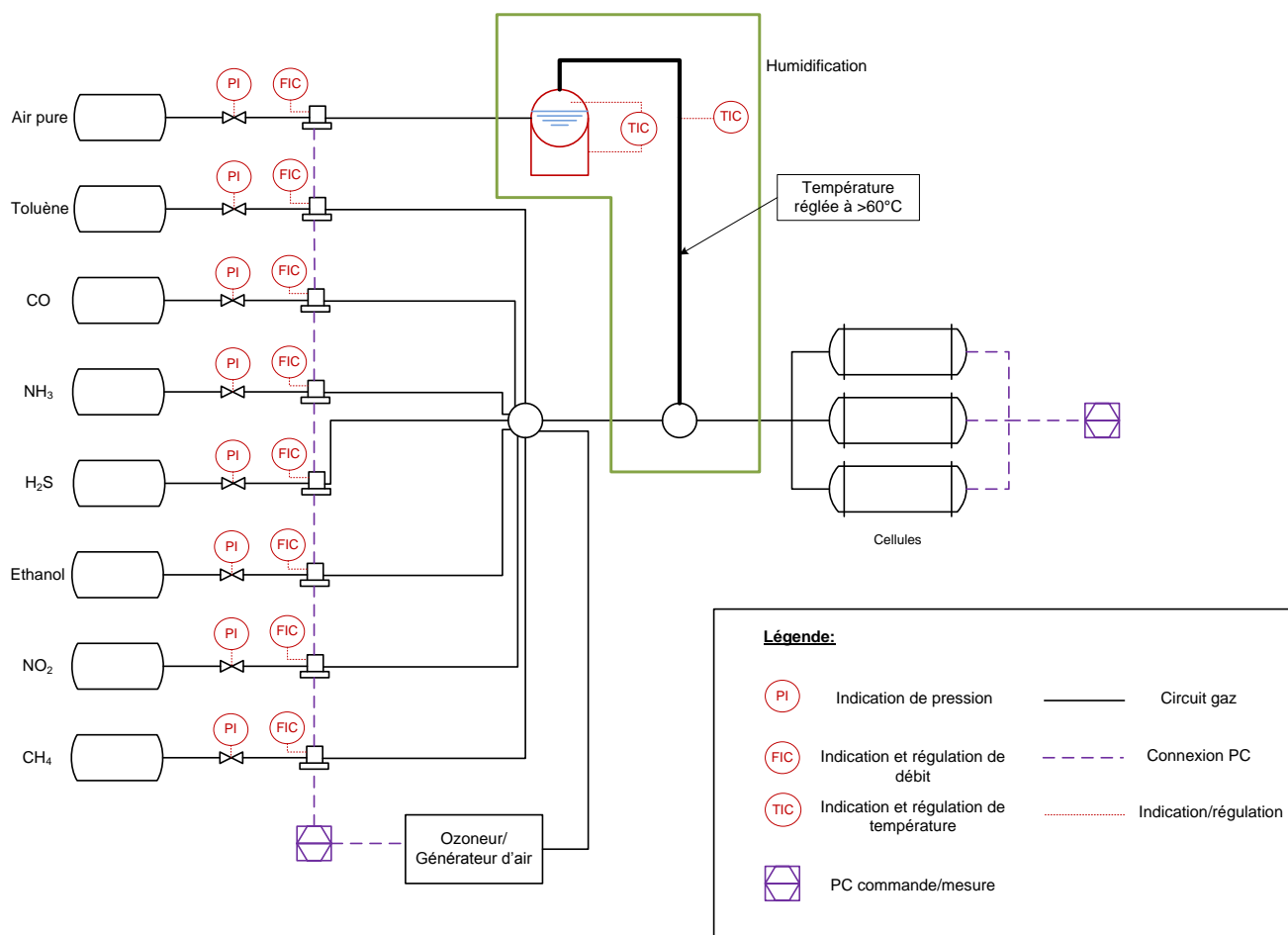


Figure 2-6 : schéma du banc de test.

Initialement, les connections gazeuses dans le banc de test étaient faites principalement avec de l'inox. Nous avons alors observé des phénomènes d'adsorption des gaz oxydants (Ozone et NO_x) ne permettant pas de générer des concentrations fiables surtout pour les faibles concentrations. Pour cette raison, la plupart des connections gazeuses ont été réalisées avec du matériel à base de PTFE (Polytétrafluoroéthylène). Pour assurer l'acquisition de données, chaque cellule est reliée à un PC via des interfaces (National Instruments USB-6212 ou USB-6008) afin d'acquérir les tensions de sortie des capteurs. A partir des tensions de charge (V_{RL}) il est facile de calculer la conductance des capteurs en utilisant les équations (1) et (2).

3.1.1. Génération des concentrations gazeuses de polluants

Le banc de test ainsi développé est capable de générer des concentrations relativement précises des polluants choisis. Cette fonction est assurée par des régulateurs de débits massiques (RDM) reliés à une source de gaz (bouteille ou instrument de génération de gaz), afin de régler les concentrations en polluants dans les cellules. Ces RDM sont pilotés par un PC à l'aide d'un programme développé sous Labview, ce dernier effectue le calcul et la gestion des débits de gaz à envoyer. Ces sources de gaz contiennent des mélanges synthétiques soit fournis avec leurs fiches de calibrage des constructeurs (par exemple Air Liquide) soit préparés par pesée au sein du laboratoire, excepté l'ozone, trop instable dans une bouteille, qui est donc généré par un ozoneur (SONIMIX 3001) capable de délivrer une concentration variable d' O_3 (entre 0 et 500 ppb) dans un débit constant de 130 L/h d'air sec et pur. Cet appareil a en effet aussi pour fonction de purifier l'air ambiant aspiré grâce à un système de cartouches filtrantes. Il sert donc aussi de générateur d'« air zéro » qui sera utilisé comme air de référence. Ce gaz (air pur + éventuel ozone) peut alors diluer les autres gaz polluants pour générer de faibles concentrations.

Les gammes de détection spécifiées par les constructeurs des capteurs (voir tableau 2-2) sont en général beaucoup plus élevées que les concentrations de polluants mesurées couramment dans l'atmosphère. Il était donc nécessaire pour nous de tester ces capteurs avec de basses concentrations afin de se rapprocher des conditions réelles d'utilisation. C'est pourquoi les faibles concentrations générées par le banc de test doivent être du même ordre de grandeur que celles rencontrées dans une pollution atmosphérique. Le programme Labview pilotant les RDM effectue le calcul des débits des gaz à envoyer. Ce programme utilise les concentrations initiales des gaz dans les bouteilles pour calculer les débits demandés selon l'équation suivante :

$$D_g = \frac{D_t * C}{C_m} \quad (8)$$

Avec :

- D_g le débit d'un gaz à envoyer via un débitmètre en L/h .
- D_t le débit de gaz total dans la cellule en L/h .
- C la concentration commandée du gaz à envoyer en ppm .
- C_m la concentration initiale du gaz dans sa bouteille en ppm .

Les incertitudes sur les concentrations générées ont pour origine d'une part les incertitudes sur les débits (connues à partir des spécifications du constructeur de RDM), d'autre part les incertitudes sur les concentrations mères, et sont calculées à partir de la loi de propagation des écart-types.

Nous avons choisi comme concentration minimale non nulle générable par notre système, la concentration pour laquelle l'incertitude relative reste inférieure à environ 20%, à l'image de la limite de quantification pour les systèmes de détection. Le tableau 2-3 suivant reprend les polluants manipulés par le banc de test tout en précisant le débit maximal associé à chaque polluant, la concentration initiale dans leurs sources, la concentration maximale

envoyée par le banc et éventuellement la concentration minimale non nulle envoyée (pour un débit de 130 L/h).

Tableau 2-3 : gamme de concentration en gaz du banc de test

Gaz	Débit maximal associé (L/h)	Concentration initiale dans la source (ppm)	Concentration min (non nulle) envoyée (ppm)	Concentration maxi envoyée (ppm)
CO	0,2	10000	$1,08 \pm 0,22$	$15 \pm 0,37$
NH ₃	10	9,4	$0,05 \pm 0,01$	$0,72 \pm 0,02$
Toluène	0,2	229,3	$0,02 \pm 0,004$	$0,35 \pm 0,01$
NO ₂	5	100	$0,27 \pm 0,05$	$3,7 \pm 0,09$
CH ₄	10	1000	$5,4 \pm 1,08$	$71 \pm 1,63$
Ethanol	10	100	$0,54 \pm 0,11$	$3,7 \pm 0,01$
H ₂ S	10	5,8	$0,03 \pm 0,01$	$0,41 \pm 0,02$
O ₃	-	-	-	0,5

3.1.2. Système d'humidification du mélange gazeux à l'aide du banc de test

Nous avons incorporé dans le banc de test, une partie dédiée à la génération d'un débit d'air humide pour contrôler le taux d'humidité relative dans les différentes cellules. La méthode d'humidification utilisée est dite « à sec » (même système d'humidification que celui utilisé par M. Kamionka) (Kamionka, 2005). Le but est d'assurer la présence d'une humidité absolue constante dans nos cellules en éliminant la possibilité de contact entre la phase liquide de l'eau et les polluants susceptibles de se dissoudre dans l'eau au cours de l'humidification du débit régulé pour obtenir la concentration désirée (voir figure 2-7). Pour effectuer une telle humidification, un faible débit (0 à 10 L/h) d'air sec est envoyé par bullage dans de l'eau pure chauffée dans un chauffe-ballon et réglée à une température T précise (par exemple $T=60\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$). L'air humide sortant de ce ballon est ainsi à la pression de vapeur saturante à T, il sera alors saturé en vapeur d'eau. Il est alors mélangé avec l'air sec contenant les polluants générés par des régulateurs de débits massiques (RDM). L'air humide ainsi que le lieu du mélange avec l'air pollué sec sont maintenus à une température supérieure à T à l'aide de cordons chauffants régulés en température pour éliminer le risque d'une condensation de la vapeur d'eau. La commande du débit d'air à humidifier est faite aussi par le même programme

Labview qui commande les RDM. Le calcul du débit d'air à humidifier se fait en se basant sur le calcul de la pression saturante de l'eau à partir de la formule de Rankine (7). Le débit d'air à humidifier est donc :

$$D_f = \frac{D_h}{H_r} = \frac{D_t * C_{eau}}{\frac{P_{sat}}{P_{atm}} * 100} \quad (9)$$

Avec :

- D_f , D_h , H_r : respectivement débit d'air à humidifier, débit de vapeur d'eau généré, humidité relative exigée en L/h .
- D_t : le débit de gaz total.
- C_{eau} : concentration d'eau souhaitée dans la cellule en g/m^3 (humidité absolue).
- P_{sat} : pression saturante de l'eau calculée à partir de la formule de Rankine, en bar .
- P_{atm} : pression atmosphérique en bar .

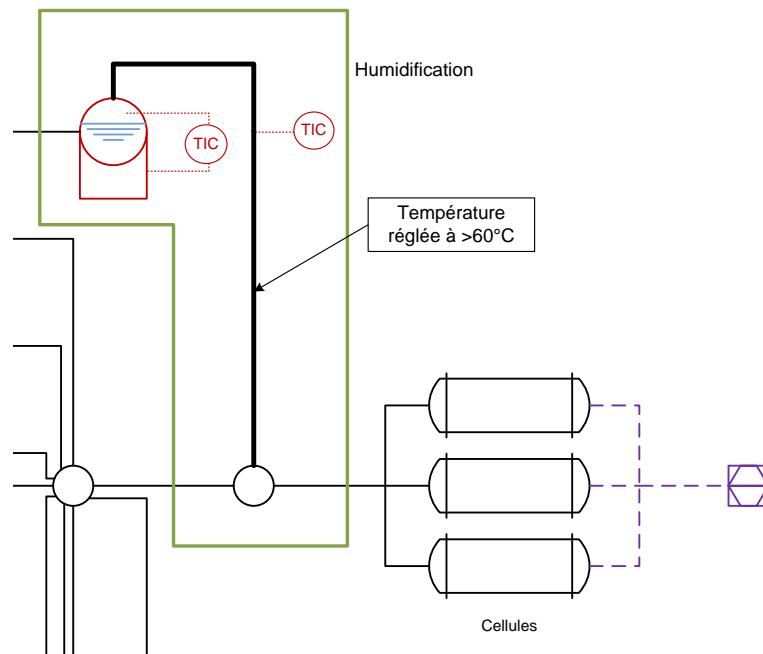


Figure 2-7 : système d'humidification séparée de l'air.

Le débit principal étant de 130 L/h (débit d'air pur et sec provenant de l'ozonneur, débit non réglable et à température ambiante), à ce débit s'ajoute le débit d'air chaud et humide de

maximum 10 L/h et un débit de source des polluants de maximum 10 L/h à température ambiante. Avec la baisse de température du débit d'air humide, nous pouvons ainsi aboutir à n'importe quelle valeur d'humidité relative désirée. Au total le débit envoyé vers nos cellules varie donc entre 130 L/h, pour un air sec pur, et 150L/h.

3.2. Banc de test développé pour les mesures « terrain »

Parallèlement à ce banc de test en laboratoire, nous avons également élaboré un banc de test adapté aux contraintes « terrain », afin d'utiliser les cellules de mesures contenant les capteurs dans des situations réelles de pollution. Comme les campagnes de mesure « terrain » ont été programmées en collaboration avec les associations de surveillance de la qualité de l'air, nous savions d'emblée que nous pourrions placer nos réseaux de capteurs à l'intérieur des stations de surveillance de la pollution atmosphérique, ce qui nous a permis de nous affranchir de l'influence des effets de la variation de la température extérieure. En effet, à l'intérieur de ces stations, la température est régulée à 23°C, et nous supposons que l'air circulant au niveau des capteurs sera sensiblement à cette température. Une bouteille d'air zéro « Air Liquide » sera utilisée sur place pour l'opération de correction périodique de dérive des capteurs qui sera détaillée dans la partie 5 de ce chapitre. La figure 2-8 présente le schéma de ce banc de test.

Par ailleurs, un système électronique a été développé afin de réaliser l'acquisition des réponses des capteurs d'une cellule de manière autonome et de les enregistrer sur une carte mémoire de type SD. Les cartes SD utilisées ont une capacité mémoire de 2 GO chacune ce qui, même avec une fréquence de mesure de l'ordre de la minute, permet d'avoir une autonomie de quelques années. Ce système fournit également une alimentation stabilisée aux capteurs. Trois cartes ont été utilisées pour les trois cellules multi-capteurs. Ce système est montré dans l'annexe A page 193.

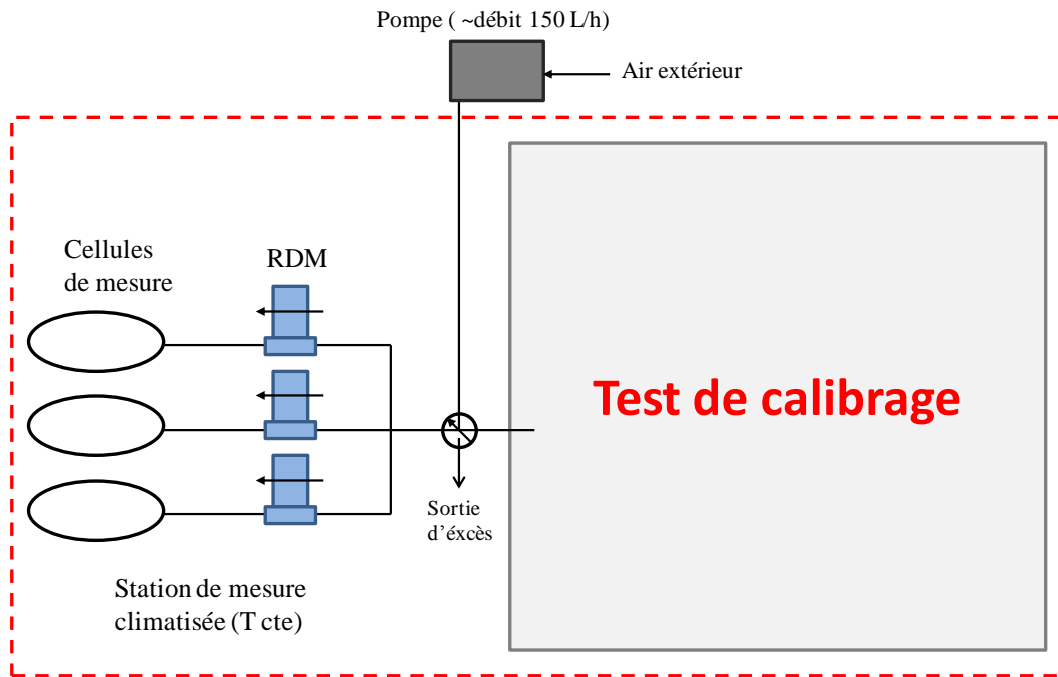


Figure 2-8 : Schéma du système de mesure à l'intérieur d'une station de surveillance de la qualité d'air

4. Réponses des capteurs

Afin de vérifier le bon comportement des capteurs dans des situations réelles, nous avons conduit un certain nombre de tests sous différentes conditions pour évaluer leur comportement.

4.1. Mesure de la conductance des capteurs

La conductance d'un capteur peut s'écrire :

$G = G_0 + g(C)$, où G_0 est la conductance avec l'atmosphère de référence, sans polluants.

G_0 est appelé « ligne de base ».

Afin de nous affranchir des écarts de ligne de base entre les différents capteurs, nous avons choisi comme variable mesurée une grandeur sans dimension appelée « réponse relative en conductance » :

$$X_i = \frac{G_i - G_{i0}}{G_{i0}} \quad (10)$$

Avec :

- X_i : la réponse relative en conductance du capteur numéro i .
- G_i : la conductance du capteur i en présence de pollution.
- G_{i0} : la conductance en atmosphère de référence du capteur i .

Cette équation est utilisable dans le cas où la conductance augmente en présence de gaz, ce qui est généralement le cas, notamment avec les gaz réducteurs. Dans le cas où la conductivité G diminue, sous certains gaz oxydants, l'utilisation de la même formule ne sera pas judicieuse car :

$$X_i = \frac{G_i - G_{i0}}{G_{i0}} \xrightarrow{G_i \rightarrow 0} -1$$

La variation des valeurs de X_i devient trop petite et difficilement utilisable. Il est alors préférable dans ce cas d'utiliser la réponse relative en résistance à la place de la conductance :

$$X_i = \frac{R_i - R_{i0}}{R_{i0}}$$

Avec :

- R_i : la résistance du capteur i sous atmosphère polluée. $R_i = 1/G_i$.
- R_{i0} : la résistance de référence du capteur numéro i , sous air de référence. $R_{i0} = 1/G_{i0}$.

Donc de façon générale, lors de la modélisation, la réponse relative en conductance sera utilisée pour les capteurs spécifiques aux gaz réducteurs qui sont: TGS2611, TGS2600, TGS2620, TGS2602, Mics5914, Mics5521 et Mics5525. Et la réponse relative en résistance sera utilisée pour les capteurs spécifiques aux oxydants qui sont : Mics2611 et Mics2710.

Pour une meilleure surveillance de l'évolution de la ligne de base de la réponse de ces capteurs en fonction du temps, la conductance de référence (G_0) correspondra toujours à la réponse des capteurs sous un air de référence. Nous avons ainsi décidé d'utiliser comme air de

référence l'air zéro (air pur) avec une humidité relative de 50% à 25°C, soit avec 10 g/m³ de vapeur d'eau. Ce G_0 sera ainsi calculé après exposition des capteurs à ces conditions de référence pendant une heure afin d'avoir un état stationnaire. Durant cette étude nous avons remarqué que chaque capteur possède une conductance spécifique sous air zéro (G_0) et que les temps de retour des capteurs à leurs lignes de bases sont différents. La mesure de l'évolution de G_0 sera développée dans la partie 5.

4.2. Effet de la température et de l'humidité sur les capteurs

Les tensions de chauffage des capteurs de gaz à semi-conducteurs commerciaux sont optimisées et précisées par leurs constructeurs. Le choix de ces tensions permet de porter la couche sensible du capteur à une température optimale, cette température est de l'ordre de quelques centaines de °C, mais n'est pas régulée. On peut donc supposer qu'un changement de cette température conduira à un changement du comportement du capteur. Il est connu que la pression partielle de vapeur d'eau a aussi un effet sur les capteurs, généralement dans le sens d'une variation positive de X_i (le même sens qu'un gaz réducteur) (Jaffrezic-Renault, et al., 1994).

Ces deux paramètres, la température et l'humidité de l'air, peuvent donc avoir des influences indésirables sur les signaux des capteurs. C'est pour cette raison que les constructeurs fournissent des graphiques qui montrent l'influence de l'humidité et de la température du gaz sur leurs capteurs. Prenons l'exemple du TGS2600, la figure 2-9 présente la courbe d'évolution du rapport R/R_0 , ou R_0 est la résistance dans de l'air à 20°C et avec 65% d'humidité relative, du capteur en fonction de la température pour trois valeurs d'humidité relative (35%, 65% et 95%).

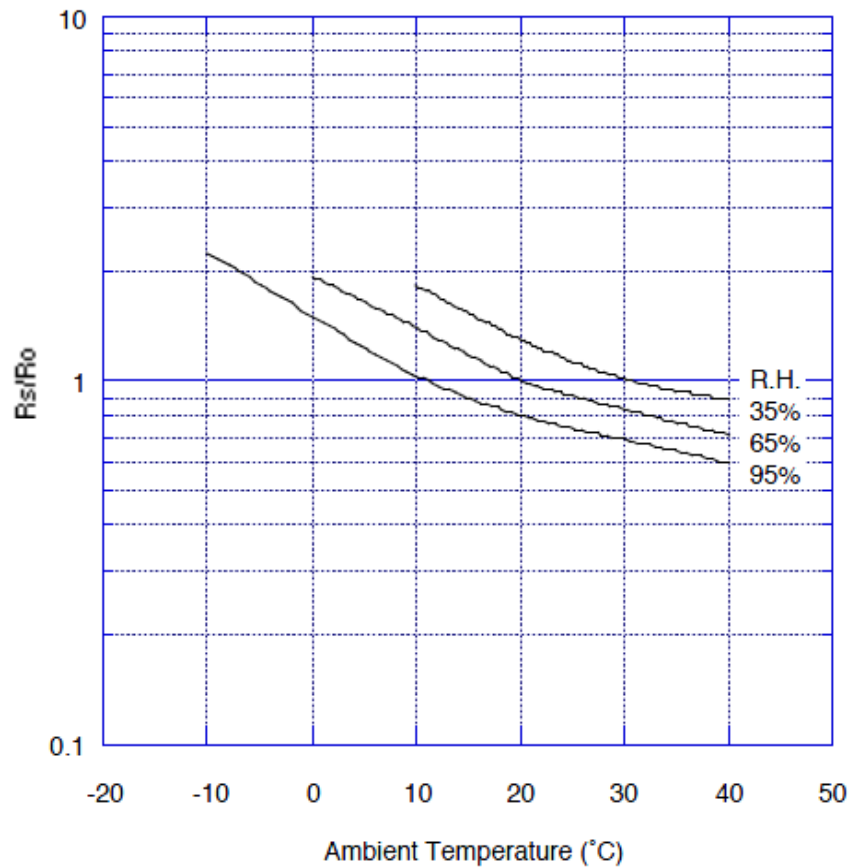


Figure 2-9 : Effet de l'humidité relative et de la température sur le signal du capteur TGS2600, données constructeur.

Pour ce capteur, la résistance de la partie sensible diminue avec l'augmentation de l'humidité relative et aussi avec l'augmentation de la température ambiante. En récupérant les données constructeurs de la figure 2-9, nous pouvons calculer la réponse des capteurs à l'humidité absolue pour différentes températures d'air (figure 2-10).

On observe alors que les 4 courbes de réponses des capteurs à l'humidité absolue pour différentes températures sont superposées, ce qui montre que ces capteurs sont sensibles à l'humidité absolue mais pas (ou peu) à la température ambiante, et que le paramètre « humidité » pertinent est l'humidité absolue, et non l'humidité relative.

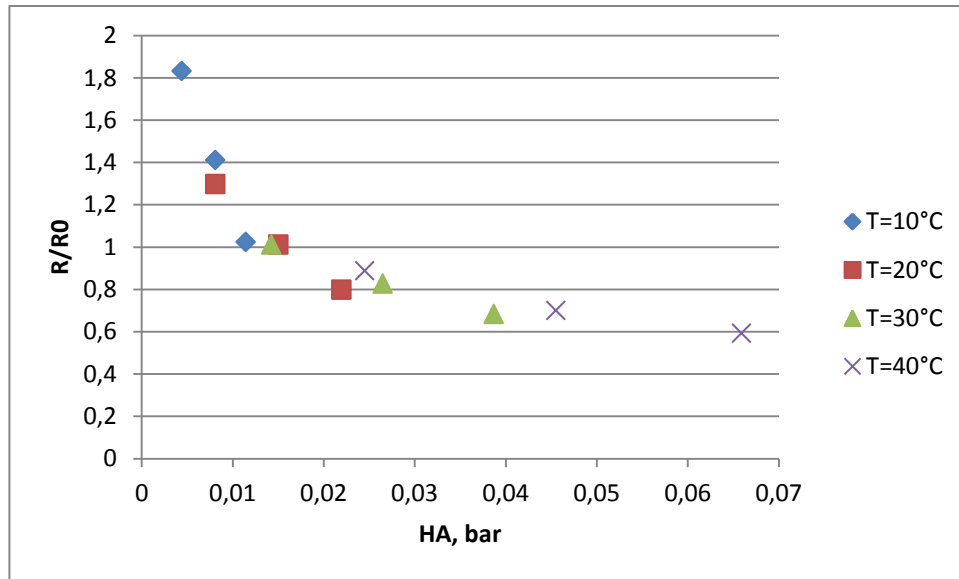


Figure 2-10 : Effet de l'humidité absolue sur le capteur TGS2600 à différentes températures calculées à partir des données du constructeur.

4.2.1. Effet de la température sur les capteurs

Afin de vérifier cette influence de l'humidité absolue du gaz sur les capteurs, nous avons effectué une mesure de la variation de la conductance des capteurs en faisant varier la température ambiante entre 0 et 40°C dans une enceinte climatique, du laboratoire de Chimie et Environnement à l'Ecole des Mines de Douai, en fixant une humidité relative constante et égale à 55%. La figure 2-11 montre la variation de la résistance relative (R_S/R_0) de la couche sensible de ces capteurs en fonction de la température, où R_0 est la résistance mesurée pour $T=25^\circ\text{C}$.

La résistance relative R_r des capteurs varie d'une façon différente d'un capteur à un autre en décroissant avec la température. Or, comme nous avons déjà montré que la température a un faible effet sur les capteurs, ces variations sont probablement dues à la variation de l'humidité absolue dans l'environnement des capteurs.

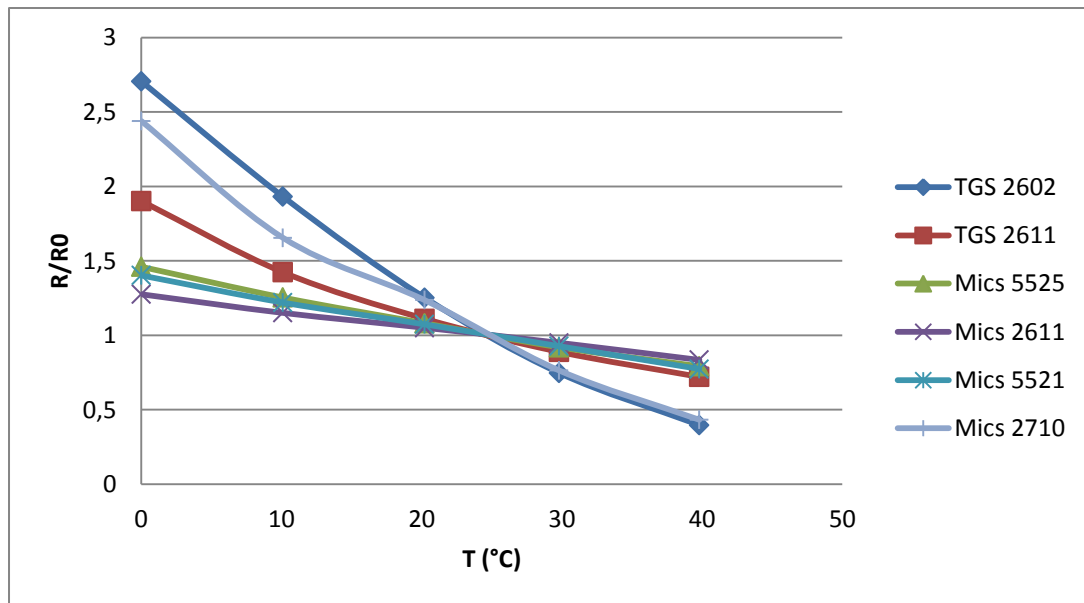


Figure 2-11 : Variation relative de la résistance des capteurs pour une variation de la température (base 1 à 23°C).

4.2.2. Influence de l'humidité absolue sur la réponse des capteurs

En effectuant le calcul permettant d'obtenir l'humidité absolue à partir de l'équation de Rankine (3), nous pouvons en déduire l'évolution de la résistance relative des capteurs en fonction de l'humidité absolue, mais à température de gaz non constante, comme illustré sur la figure 2-12.

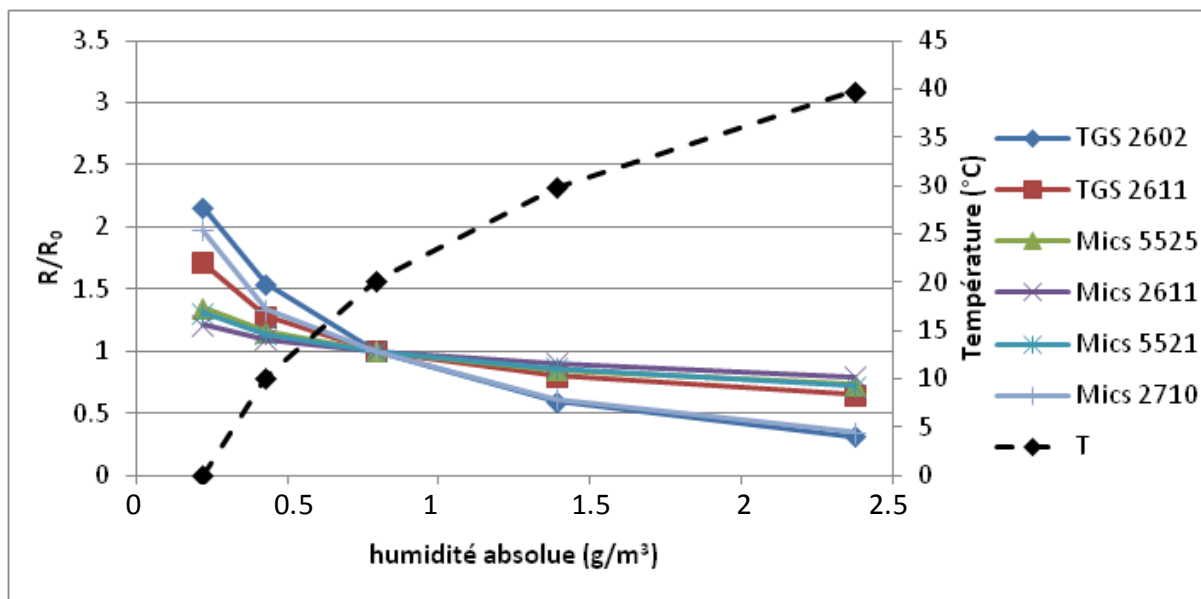


Figure 2-12 : Evolution des résistances relatives des capteurs en fonction de l'humidité absolue. La variation de la température ambiante est montrée dans la courbe pointillée.

Cette variation de R_r montré est supposée due à la variation de l'humidité absolue si nous considérons que la température de gaz a un effet négligeable sur les réponses des capteurs. L'enceinte de mesure utilisée ici nous permet hélas de ne modifier que la température à humidité relative constante et donc pas l'humidité absolue. Il est donc impossible ici de séparer les 2 effets « température du gaz » et « vapeur d'eau ».

Ce test a donc été répété dans le banc de test de l'école des mines de Saint Etienne. Dans ce banc nous pouvons contrôler la concentration en humidité absolue dans nos cellules et donc étudier la variation de la résistance relative des capteurs par rapport à l'humidité absolue, mais ici, c'est la température du gaz qui est non contrôlée et donc à la température ambiante, soit environ 25 °C. La figure 2-13 présente la variation de la résistance relative des capteurs par rapport à la variation de l'humidité absolue.

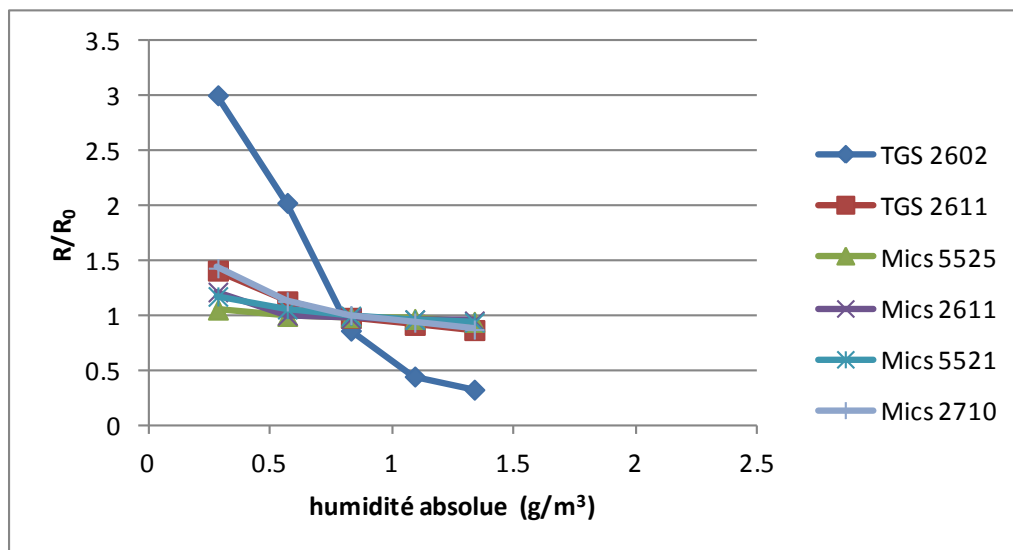


Figure 2-13 : Variation de la résistance (base 1 pour 8 g/m³) de l'humidité absolue pour $T_{\text{gaz}} = 25 \text{ °C}$.

Les deux figures 2-12 et 2-13 montrent la même forme d'évolution des résistances des capteurs avec des amplitudes semblables, sauf pour le TGS 2602 pour des causes inconnues. La différence entre les deux figures montre que si un effet « température du gaz » existe, il est non déterminant pour des variations peu importantes autour de la température ambiante. D'ailleurs A. Valleron a vérifié une absence d'effet de la température des gaz entre 20 et

200°C sur des capteurs au dioxyde d'étain, à la température de chauffage régulée (Valleron, 2011).

L'influence de la température n'ayant pas pu être déterminée avec précision, nous avons décidé de poursuivre nos expériences dans des situations où la température ambiante reste constante ou subit peu de variation pour s'affranchir de son effet, même si cette température est systématiquement mesurée et pourrait donc être intégrée dans les modèles. Par contre nous ne pouvons pas garder constante l'humidité absolue de l'air qui peut varier de façon importante suivant les conditions météorologiques.

La figure 2-14 présente la variation de la réponse relative (X_i) du capteur TGS2602 à des concentrations d'éthanol mesurées à différentes valeurs d'humidité absolue sous une température constante (25 °C). Nous remarquons comme prévu un décalage ainsi qu'une légère variation de sensibilité. Rappelons que la conductivité de référence, G_0 , correspond à 25°C et 50% d'humidité relative soit environ 1g/m³ en absolue.

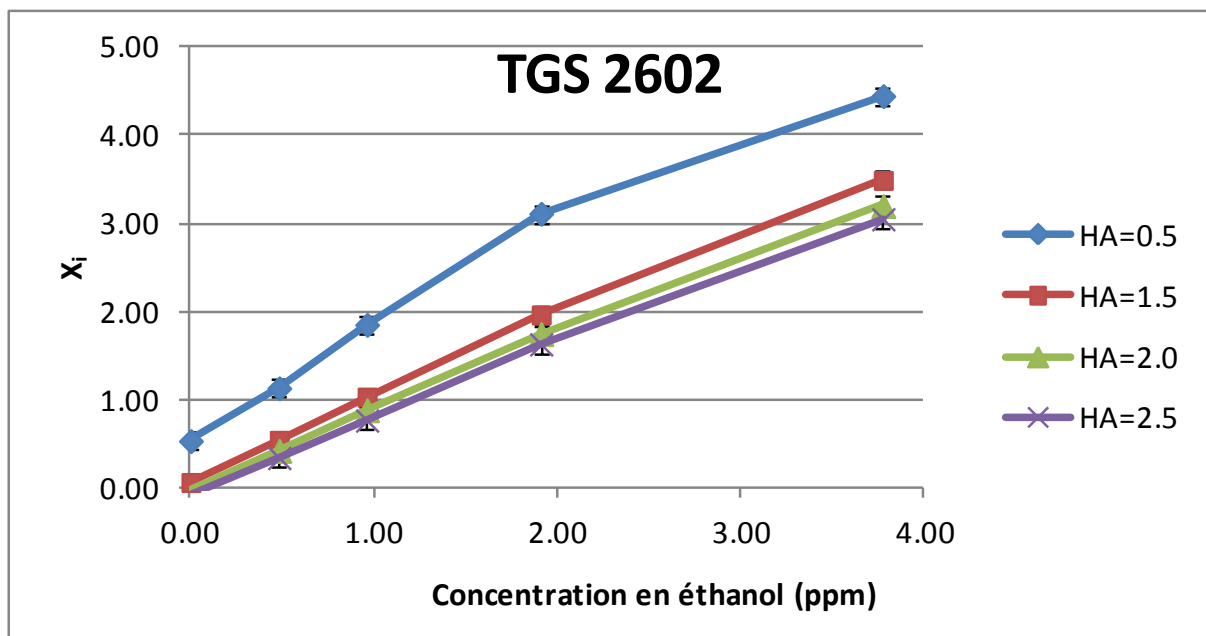


Figure 2-14 : Réponses du capteur TGS2602 pour une concentration croissante en éthanol mesuré à des humidités absolues (g/m³) différentes.

La variation de l'humidité absolue a non seulement un effet sur la variation de G_0 du capteur, mais elle a aussi une influence sur la sensibilité du capteur pour un polluant donné (pente). Il est donc indispensable de mesurer l'humidité de l'air en continu et de l'incorporer dans les modèles à développer, ce qui a été fait par la suite en considérant l'humidité comme une donnée d'entrée dans les réseaux de neurones (voir chapitre 3 partie 4.3.).

4.3. Réponse des capteurs de gaz – signature des gaz

Après avoir vérifié la présence d'un effet de l'humidité absolue de l'air sur la réponse des capteurs, nous avons poursuivi les tests des capteurs avec des polluants gazeux présents habituellement dans l'air ambiant.

4.3.1. Cas des gaz réducteurs

La réponse des capteurs vis-à-vis des gaz réducteurs générés par le banc de test de validation (éthanol, toluène, CH_4 , CO , NH_3 , H_2S) a été étudiée. Lorsqu'on injecte un gaz réducteur seul vers un capteur de gaz à semi-conducteur, on obtient une variation de conductance positive qui se traduit par une variation positive de la conductance relative (X_i). Prenons l'exemple de l'injection de l'éthanol qui est un gaz réducteur. Une injection d'une concentration croissante produit une réponse croissante monotone sur les capteurs (et ici presque linéaire, mais ce n'est pas le cas avec tous les couples gaz-capteur). La figure 2-15 illustre cette réponse sur 5 capteurs censés détecter les COVs et les odeurs : TGS2600, TGS2602, TGS2611, TGS2620, MICS5521.

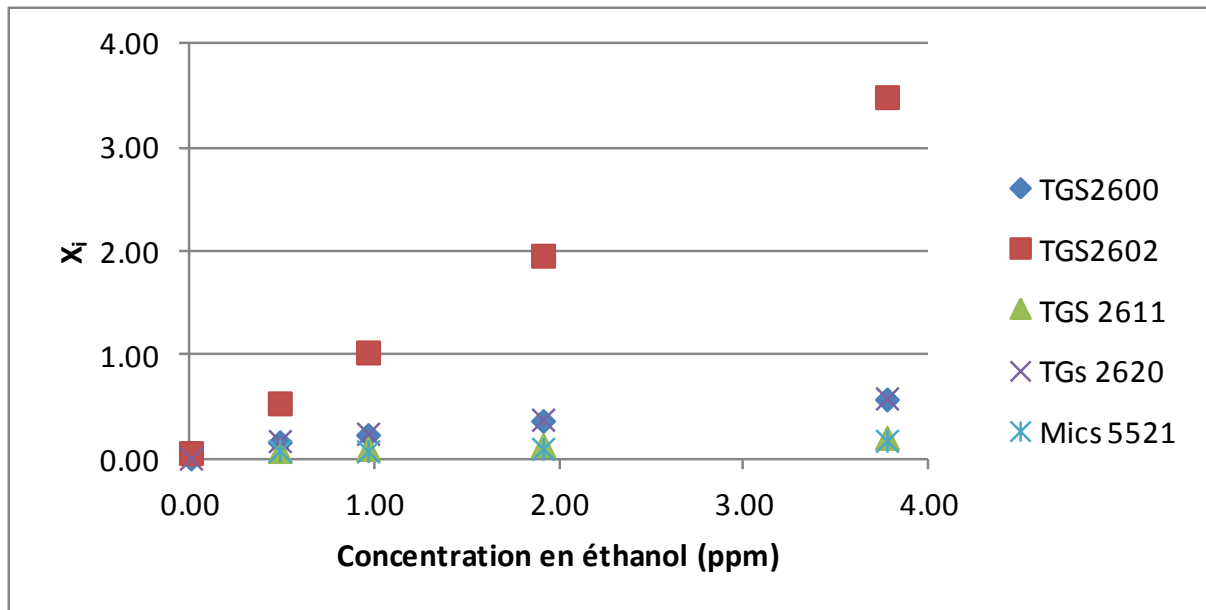


Figure 2-15 : Réponse relative en conduction des capteurs à l'éthanol

Par ailleurs, chacun de ces capteurs possède une sensibilité différente, ce qui signifie qu'ils possèdent une réponse spécifique à un gaz donné. Les capteurs vont donc présenter une "signature" spécifique pour un groupe de gaz donnés. La variété et la non colinéarité de ces signatures sont des conditions indispensables pour la détection sélective à l'aide d'un réseau de capteurs. La figure 2-16 illustre la réponse relative du réseau de capteurs pour trois gaz réducteurs (toluène, éthanol, ammoniac) pour des concentrations spécifiques. Rappelons que la non réponse se traduit par une réponse relative nulle.

Les signatures sont bien spécifiques de chaque gaz. Ainsi, le toluène est détecté uniquement par le capteur TGS 2602, et la différence entre la forme de la signature l'éthanol et celle du NH_3 n'est pas seulement due aux différences de concentrations de ces gaz, mais bien aux écarts de sensibilité de chaque capteur.

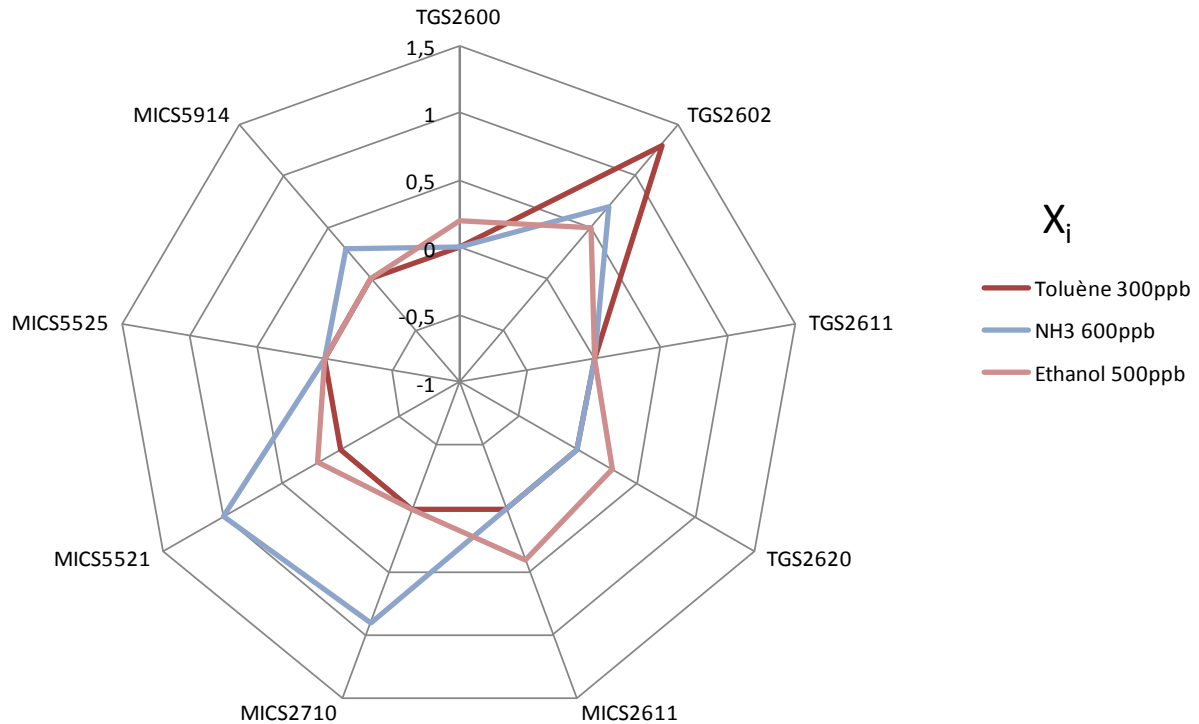


Figure 2-16 : réponse relative des capteurs (X_i) pour des gaz réducteurs.

4.3.2. Cas des gaz oxydants

Les gaz oxydants les plus rencontrés dans une pollution atmosphérique sont le NO_2 et l' O_3 . La figure 2-17 montre que seuls 2 capteurs répondent à ces gaz : le Mics2710 et le Mics2611.

Comme prévu, ces capteurs répondent négativement en raison des propriétés oxydantes de ces gaz. Nous remarquons que le capteur Mics2710 initialement prévu pour détecter uniquement le NO_2 présente une réponse de -1 sous une faible concentration d' O_3 , ce qui signifie un état de très grande résistance du capteur. Comme indiqué dans la partie 1.2, l'utilisation de la « réponse relative en conductance » n'est pas appropriée pour ce type de capteur (saturation de la variable X_i à -1), nous la gardons ici dans un but de comparaison. Dans les futures utilisations sur terrain, les réponses du Mics2710 et du Mics2611 seront donc présentées en variation relative de résistance pour mieux suivre les variations de la présence de NO_2 et O_3 dans l'atmosphère, afin d'éviter ce problème de saturation du signal à -1.

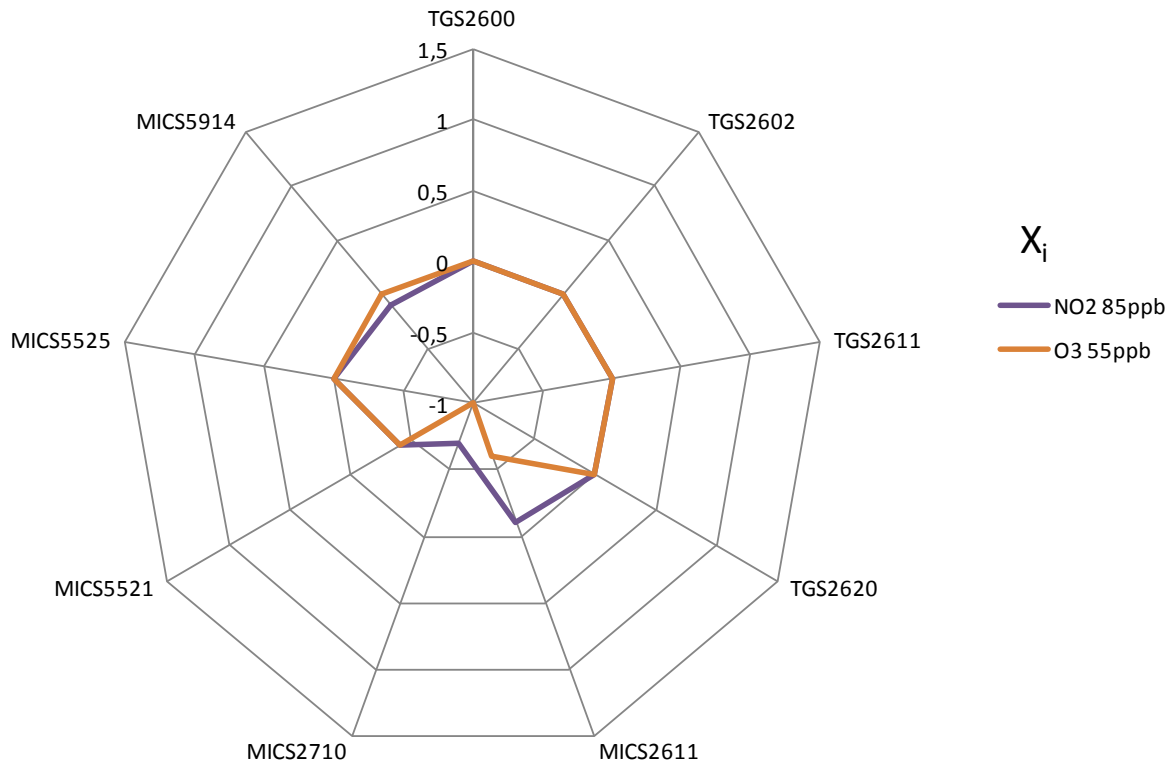


Figure 2-17 : Réponse relative (X_i) des capteurs pour des gaz oxydant (O_3 et NO_2).

4.3.3. Signature des différents gaz sur les différents capteurs

Au total, les capteurs ont été testés avec 7 polluants gazeux différents. Les X_i obtenues sont données en figure 2-18 ; les signatures observées pour chaque gaz différent bien les unes des autres sur l'ensemble des capteurs. Les signatures observées pour les différentes concentrations d'un même gaz diffèrent aussi les unes des autres comme montré dans la figure 2-19 en gardant toujours des formes semblables.

Nous remarquons sur ces figures que les capteurs ont des sensibilités encore importantes pour des concentrations sensiblement en deçà de celles préconisées par les constructeurs, ce qui nous permet d'avoir de bon espoir pour la pollution atmosphérique.

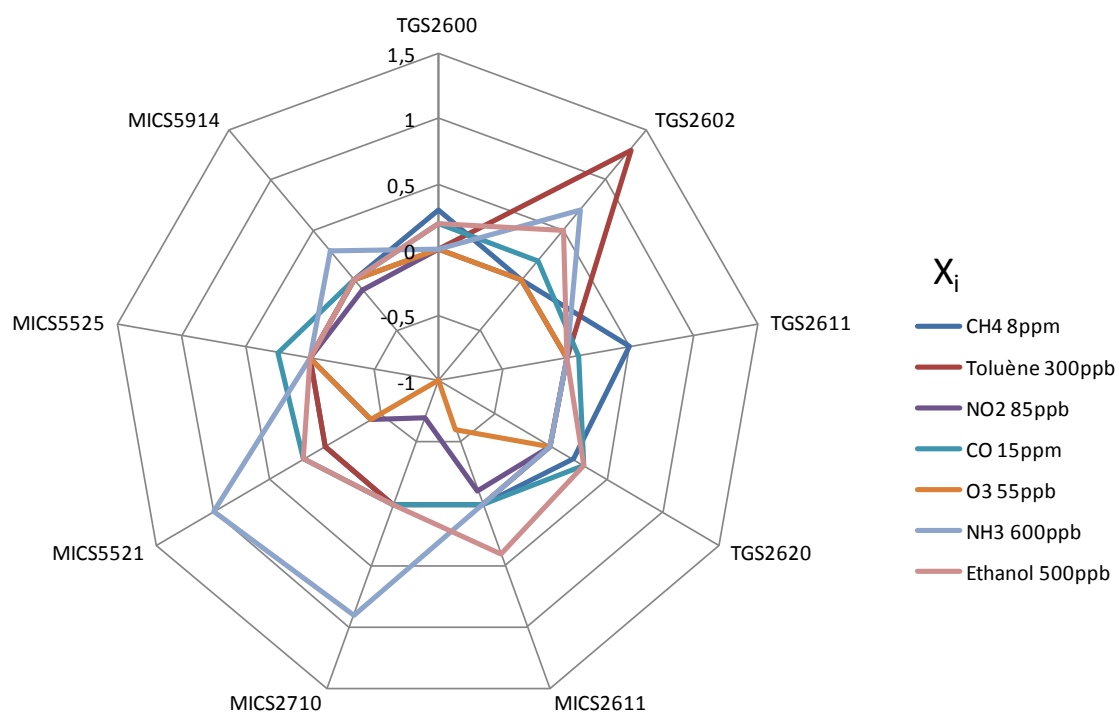


Figure 2-18 : Réponse relative (X_i) des capteurs pour des différents gaz.

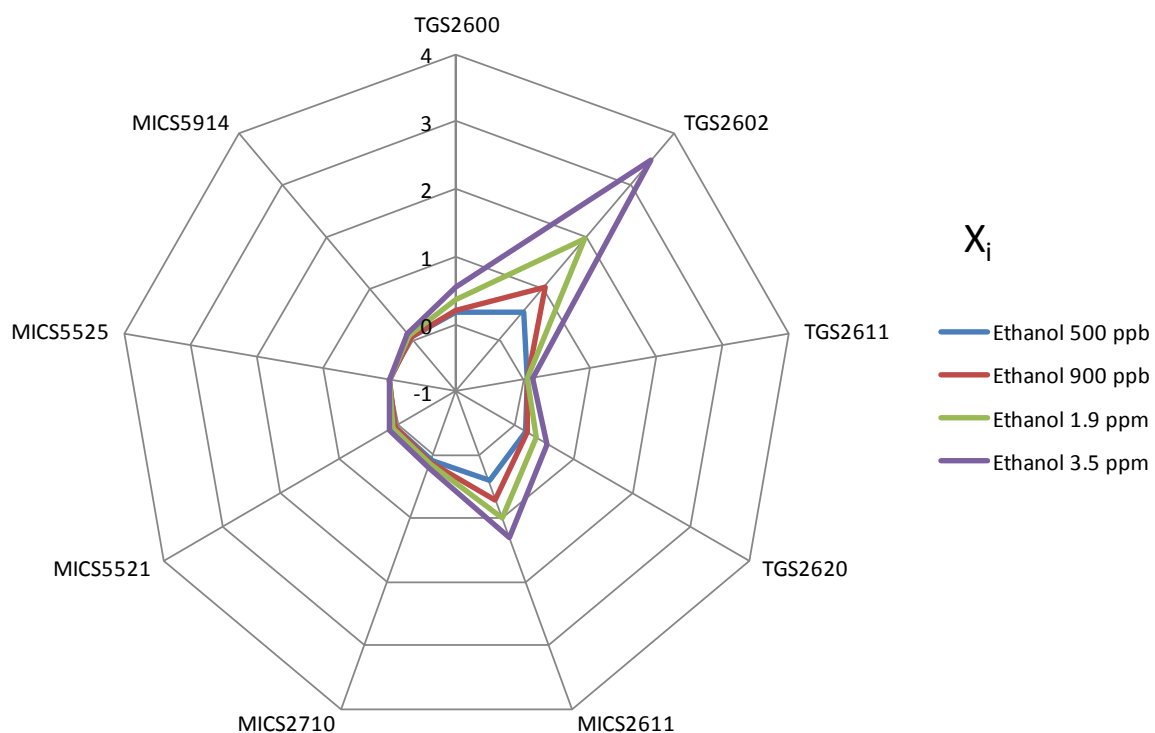


Figure 2-19 : Réponse relative (X_i) des capteurs pour des différentes concentrations d'éthanol.

Il est donc peut-être possible à partir de ces signatures, d'établir un modèle capable de prédire les concentrations avec les gaz testés : c'est l'approche quantitative. Cette approche est limitée par le nombre de gaz à détecter et le nombre de capteurs dans un réseau. Un grand

nombre de gaz à détecter nécessite un grand nombre de capteurs et un modèle de plus en plus complexe. C'est pourquoi nous avons décidé d'adopter une approche plus qualitative dans nos futures études.

5. Gestion des écarts de sensibilité et de la dérive de la ligne de base des capteurs

Pour le bon fonctionnement de ces capteurs, les constructeurs spécifient une période de « maturation ». En général, cette période est de 48h. C'est une période durant laquelle le capteur neuf doit être alimenté en courant électrique de sorte à atteindre sa sensibilité optimale d'une part et une certaine stabilité temporelle d'autre part. Après maturation, chaque nouvelle utilisation des capteur nécessite encore une période de stabilisation après mise sous tension pour qu'ils atteignent leurs sensibilité optimale et aboutissent à leur état stationnaire décrit par la ligne de base G_0 , mesurée sous air de référence. Pour rappel, l'air de référence est issu d'une bouteille d'air zéro (air pur) avec une humidité relative de 50% à 25°C. La figure 2-20 montre le temps nécessaire pour que les capteurs atteignent leur ligne de base.

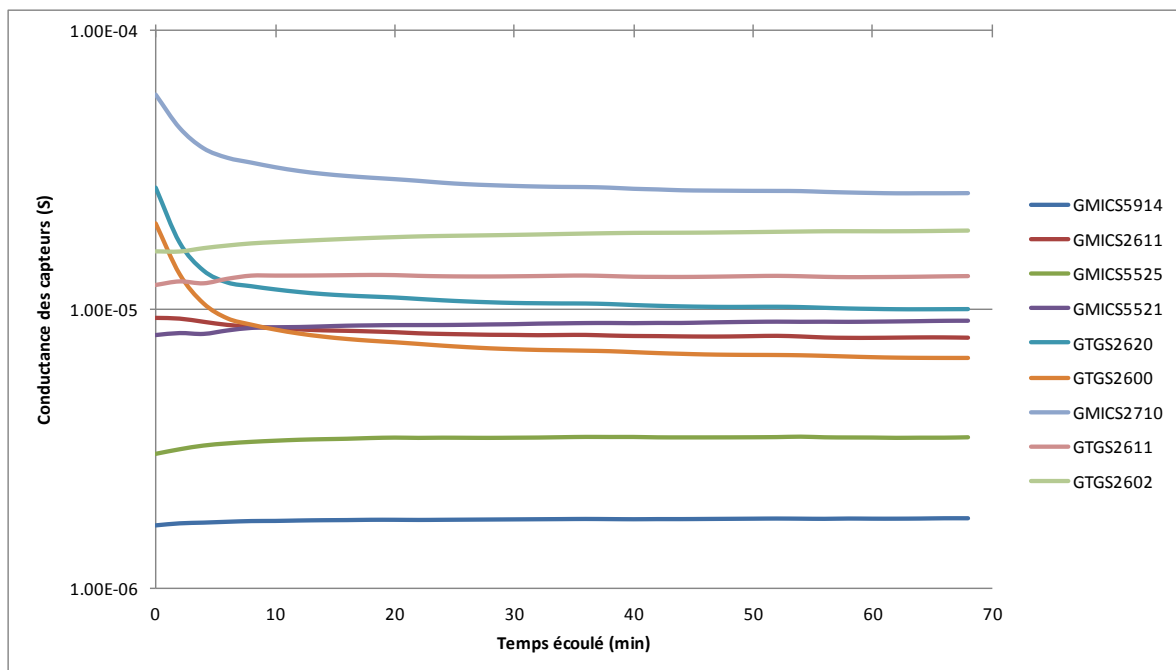


Figure 2-20 : Stabilisation des réponses des capteurs « matures » au cours de la période de préchauffage.

Nous remarquons que ce temps n'est pas le même pour tous les capteurs, il peut varier de quelque minutes jusqu'à quelques dizaines des minutes selon la nature des capteurs. Nous constatons qu'au delà d'une période de 30 min, tous les capteurs aboutissent à un état considéré comme stationnaire. Nous avons donc décidé de débiter nos expériences par une période d'une heure de recouvrement sous air zéro humide (25 °C ; 50% HR). que tous les capteurs atteignent correctement leurs lignes de base.

5.1. Evolution à long terme des lignes de base

Durant la période de développement et d'étalonnage de notre cellule prototype (cellule N° 1), nous avons étudié la dérive à long terme des lignes de base des capteurs sous atmosphère contrôlée (25 °C ; 50% HR), en mesurant chaque mois cette conductance de référence pour chaque capteur air de référence. Durant une période de 7 mois (Février 2010 – Aout 2010) nous avons remarqué une évolution non négligeable des conductances de référence (G_0) des capteurs. La figure 2-21 illustre l'évolution du X_i subit entre le mois de février et août 2010 sur la cellule prototype.

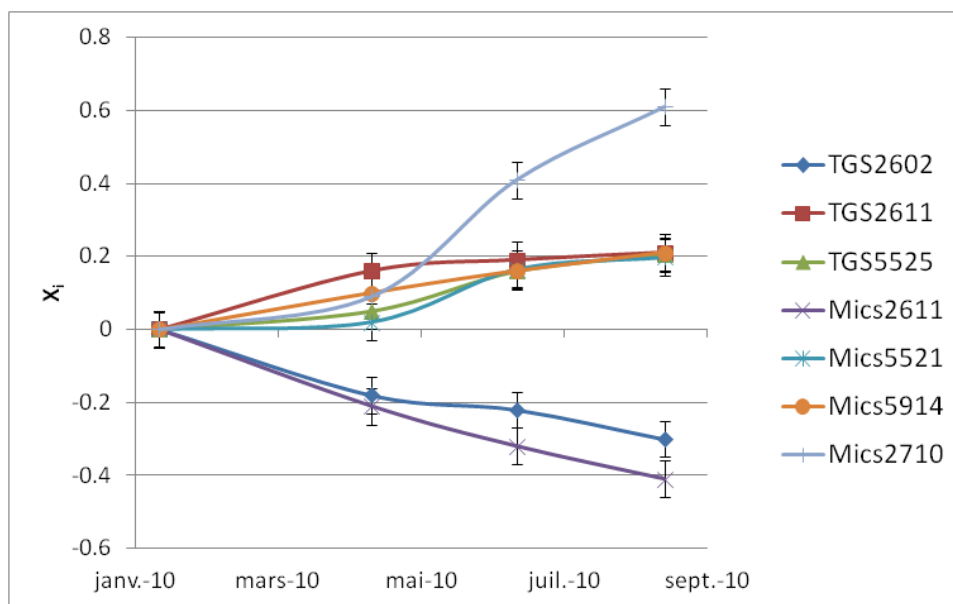


Figure 2-21 : Evolution de la ligne de base des capteurs durant 7 mois d'utilisations.

5.2. Test de calibrage de G_0

Cette dérive non négligeable pourra entraîner des problèmes de validité temporelle des modèles réalisés. Ainsi, une mesure périodique de cette évolution est nécessaire afin d'assurer le suivi de la ligne de base des trois cellules de mesures et permettre de corriger son éventuelle dérive. Il est impératif que ce test soit toujours répété sous les mêmes conditions de référence (25 °C ; 50% HR). Sur le « terrain », ce test a été pratiqué avec une fréquence hebdomadaire à l'aide d'une bouteille d'air zéro comme montré dans la figure 2-22. Le débit d'air zéro provenant de cette bouteille est contrôlé à l'aide de deux RDM pour assurer un débit d'air humide à 50% HR en envoyant un débit d'air sec de 75 L/h mélangé avec un débit d'air humide de 75L/h.

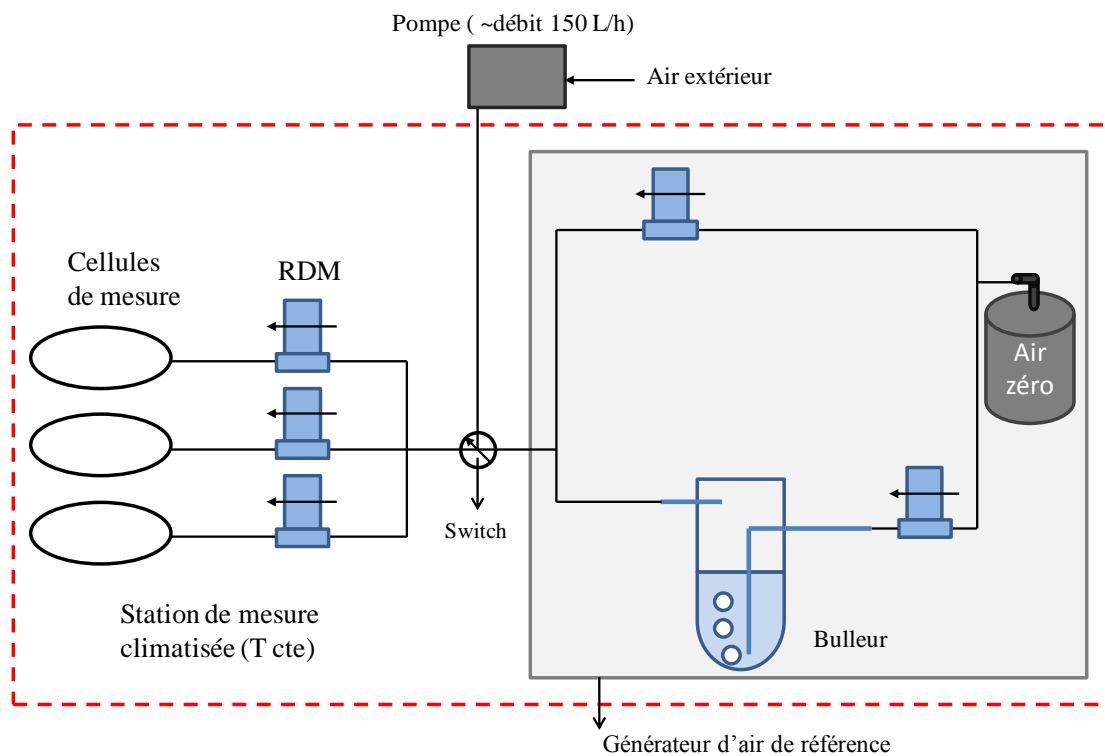


Figure 2-22 : Implémentation d'un test de suivi de ligne de base via le banc de test « terrain ».

Les dérives enregistrées diffèrent selon les capteurs, même des capteurs appartenant au même type peuvent présenter des dérives différentes ce qui est illustré sur la figure 2-23 avec

l'évolution de la ligne de base de capteurs Mics5525 issues des 3 cellules de mesure pendant une période de 6 mois.

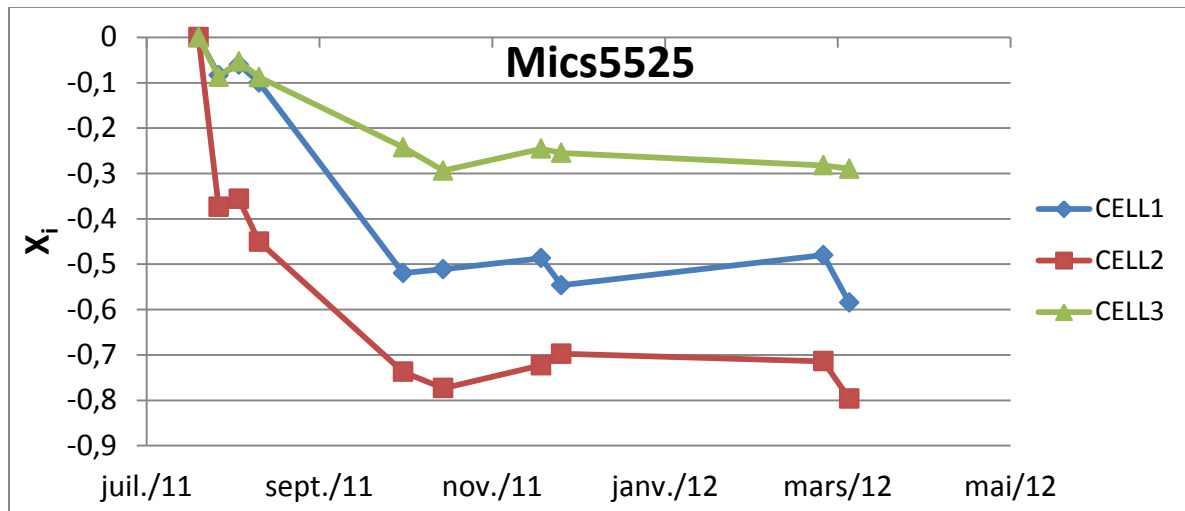


Figure 2-23 : Variation de la conductance relative des capteurs Mics5525 pendant une période de 6 mois (Juillet au mars) pour les trois cellules.

Cette figure montre une évolution décroissante quasi-similaire des conductances de référence des capteurs Mics5525. L'amplitude de la variation varie d'un capteur à un autre, la variable montrée est la réponse relative en conductance X_i (voir §4) calculée par rapport à un point fixe (premier G_0). Nous remarquons que ces capteurs ont subi une variation décroissante pendant les premiers mois (août à novembre), et après cette période leurs lignes de références se stabilisent. Ceci n'est pas le cas pour l'ensemble des capteurs, comme par exemple pour le TGS2602 illustré dans la figure 2-24.

Ces capteurs présentent des évolutions différentes des G_0 avec des variations brusques dans des sens positifs ou négatifs. Cette évolution non reproductible d'un capteur à un autre (y compris entre les capteurs de même type) a été remarquée sur presque la totalité de nos capteurs issus des différentes cellules (sauf pour le cas des Mics5525).

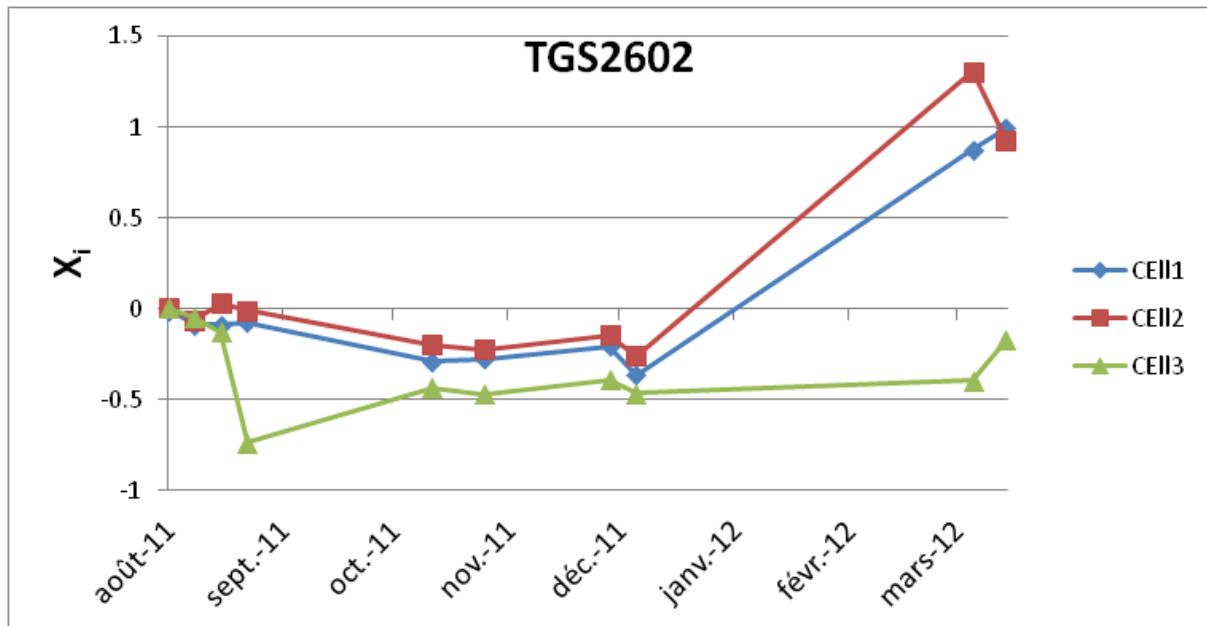


Figure 2-24 : Evolution des lignes de base pour le capteur TGS2602

Ces variations non prévisibles montrent la difficulté de modéliser mathématiquement les dérives des capteurs en fonction du temps. Toutefois, certaines variations brusques observées sur ces lignes de base et communes aux trois capteurs à la fois (fig. 2-23 et 2-24) ont probablement une cause extrinsèque aux capteurs (problème d'instabilité accidentelle dans la génération du gaz de référence).

Afin de limiter l'effet de ces variations intempestives, nous avons décidé de lisser les valeurs de G_0 en calculant la moyenne sur deux points. Cette méthode a abouti à des évolutions des lignes de base monotones et donc plus réalistes. La figure 2-25 donne les évolutions des G_0 lissées pour la cellule n° 3.

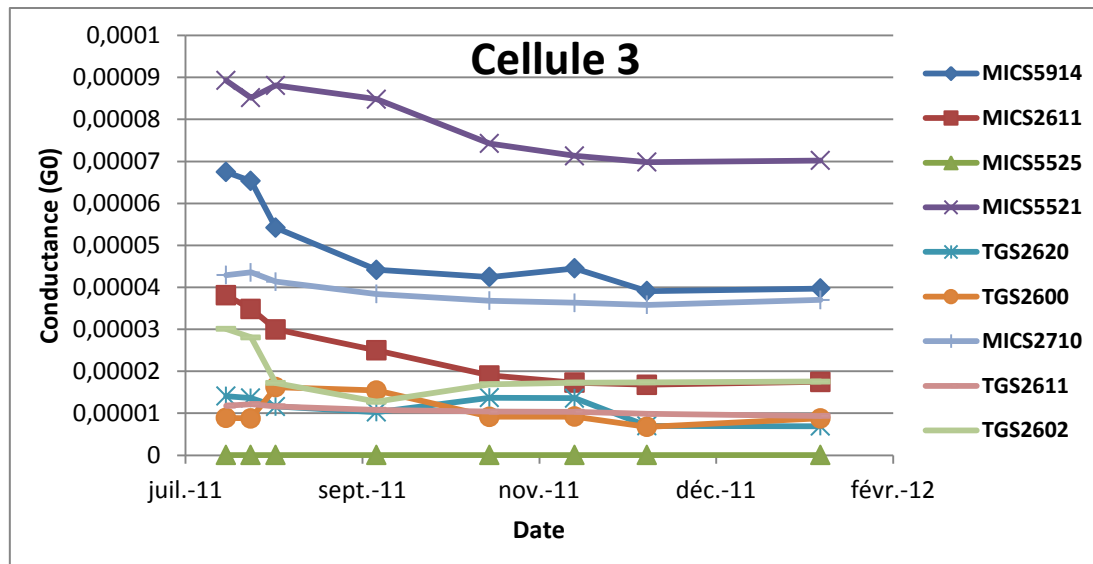


Figure 2-25 : évolution des conductances relatives lissées.

5.3. Implémentation d'une standardisation de sensibilité

Un des problèmes des capteurs de gaz à semi-conducteurs est la non répétabilité de leurs réponses entre individus d'un même type. Nous avons vu dans le paragraphe précédent qu'il était possible de faire abstraction de la ligne de base, il reste à traiter les différences de sensibilité. L'injection simultanée de gaz dans les 3 cellules, permet de comparer ces écarts de sensibilité. Ceci est illustré dans la figure 2- 26 avec les réponses relatives en conductance X_i des 3 capteurs de type TGS2600 provenant de 3 cellules différentes sous des concentrations croissantes en méthane.

Ces capteurs de même type montrent une évolution monotone avec la croissance de la concentration en méthane. Par contre la pente de cette évolution varie d'une cellule à l'autre. Il s'agit du problème de non répétabilité des sensibilités. Un tel problème risque d'induire des signatures différentes sur chacune de nos cellules pourtant exposées à un même gaz. Afin de pallier ce problème, nous proposons une démarche de standardisation des différentes sensibilités des capteurs.

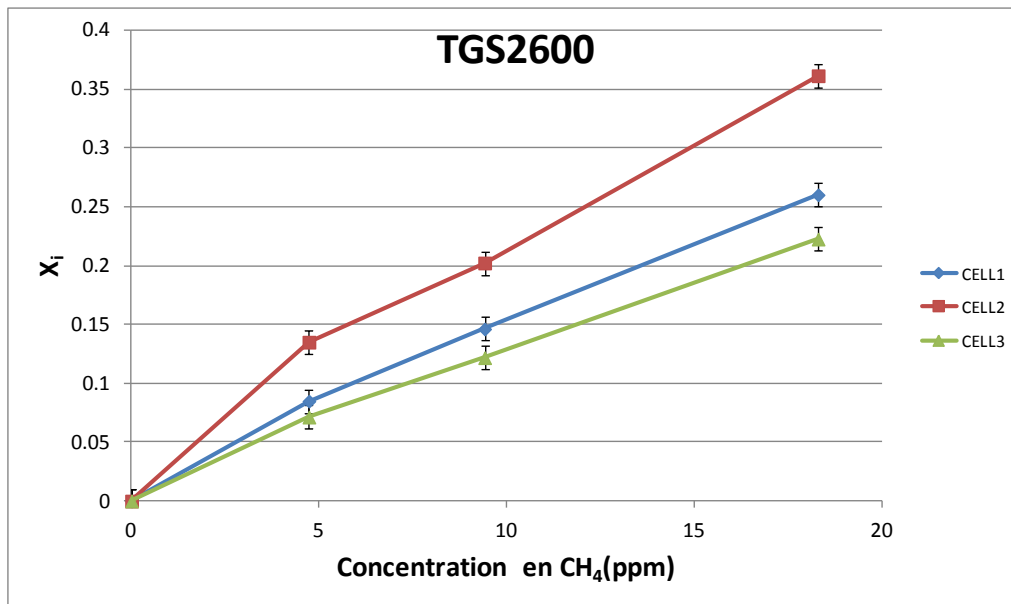


Figure 2-26 : Problème de répétabilité entre capteurs de même type, différentes sensibilités sous une concentration croissante en méthane.

Le but de cette démarche est d'homogénéiser le mieux possible les réponses entre les cellules en normalisant les sensibilités des capteurs de mêmes types. Dans cette démarche la variation de la conductance relative d'un capteur (X_i) (ou la variation de la résistance relative si il s'agit d'un capteur pour gaz oxydants) doit être standardisée en fonction d'un gaz spécifique pour chaque type de capteur étudié. X_i sera divisée par la réponse relative du capteur « i » calculée sous une concentration choisie de son gaz de standardisation choisi (X_{i_s}).

$$S_i = \frac{X_i}{X_{i_s}} \quad (11)$$

Quatre gaz ont été choisis pour cette démarche de standardisation, ils sont censés représenter les familles de polluants présents en pollution atmosphérique (CO, CH₄, NO₂ et NH₃). Ils seront aussi représentatifs des types de capteurs : CO pour les capteurs de CO, CH₄ pour les capteurs de CH₄, NO₂ pour les capteurs des gaz oxydants, et NH₃ pour les capteurs des odeurs et des COV (tenant compte des propriétés réductrices de ces gaz). Nous montrons

dans le tableau 2-4 suivant les différents types de capteurs utilisés dans les cellules que nous avons développé avec leurs gaz de standardisation choisis et leurs concentrations associées.

Tableau 2-4: Les différents capteurs de gaz utilisés avec leurs gaz cibles et leurs gaz de standardisations accordées

Capteurs	Gaz cibles définis par le constructeur	Gaz de standardisations et leurs concentrations choisies	
MICS5525	CO	CO	20 ppm
MICS5521	CO, COV, Hydrocarbures	CO	20 ppm
MICS2710	NO ₂	NO ₂	0,4 ppm
MICS2611	O ₃	NO ₂	0,4 ppm
TGS2600	Contaminants généraux de l'air	NH ₃	4 ppm
TGS2602	COV, Odeurs, H ₂ S.	NH ₃	4 ppm
TGS2620	Alcools et vapeurs solvants	NH ₃	4 ppm
MICS5914	NH ₃	NH ₃	4 ppm
TGS2611	Gaz combustibles (Méthane, Butane...)	CH ₄	20 ppm

Ce test est effectué dans des conditions de référence de température et d'humidité (25 °C, 50% HR) pour assurer l'homogénéité de différentes standardisations. Il est également utilisé comme un outil de diagnostic pour vérifier tout mauvais comportement de l'un des capteurs comme par exemple une perte de sensibilité à un gaz cible. Il offre par ailleurs la possibilité de changer un capteur défaillant sans remettre en cause le modèle de prédiction.

5.4. Effet de la standardisation sur les sensibilités mesurées

L'effet de cette standardisation sur nos capteurs est présenté sur la figure 2-27. Après standardisation, on note des similitudes entre les réponses.

Dans la figure 2-27, le graphe « A » représente les variations des conductances relatives des capteurs des 3 cellules sous leurs gaz cibles avant standardisation, et le graphe « B » montre ces variations après standardisation. Nous remarquons dans le graphe A des écarts importants entre les réponses des capteurs sous les différents gaz. La standardisation normalise ces réponses de sorte qu'elles soient ramenées à 1 sous leur gaz cible.

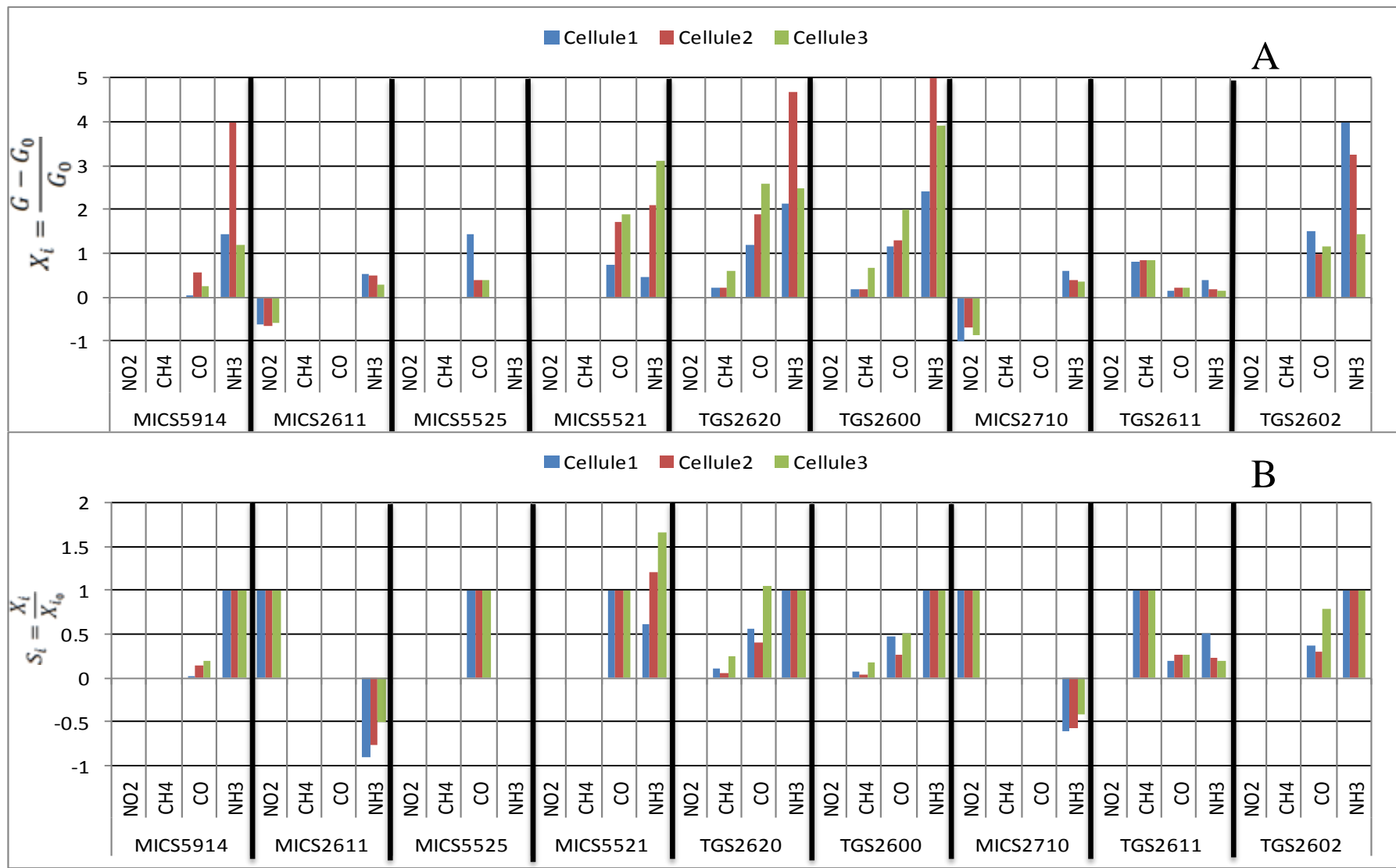


Figure 2-27 : Graphe A représente les variations des conductances relatives des capteurs calculées sous les gaz de standardisation, graphe B représente les réponses standardisées des capteurs calculées selon l'équation n°11 (test effectué en juillet 2011).

Même après la standardisation, nous notons toutefois une différence entre les réponses des capteurs pour leurs gaz non cibles, ce qui constitue la faiblesse de la méthode, car il n'y a bien sûr qu'un gaz de standardisation par type de capteur. Par contre cette méthode a diminué considérablement l'amplitude de ces variations. Ce test sera effectué trimestriellement pour vérifier l'existence éventuelle d'une dérive de sensibilité et de la corriger, ou de changer le capteur.

5.5. Dérive temporelle des sensibilités

Les capteurs semi-conducteurs subissent au-delà de la non répétabilité de leurs réponses, une dérive de sensibilité. Cette dérive a été détectée en effectuant le test de standardisation sur des périodes de temps longues. La figure 2-28 montre la variation temporelle de la sensibilité des capteurs de type TGS 2620 issues des différentes cellules sous leurs gaz cible de standardisation mesurée en juillet 2011, octobre 2011 et janvier 2012. Une décroissance importante de cette sensibilité est observée, notamment pour le méthane.

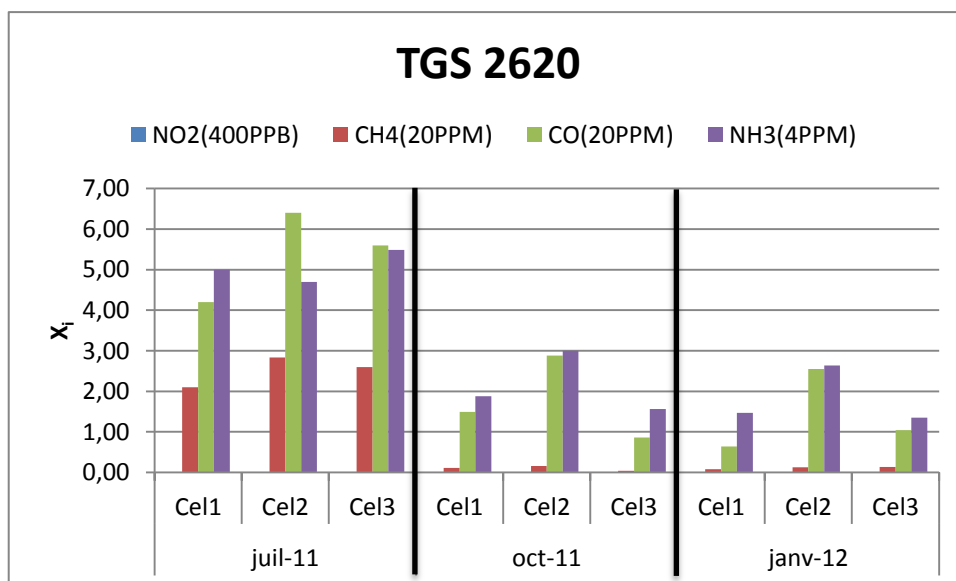


Figure 2-28 : évolution temporelle de la réponse non standardisée (X_i) du TGS2620.

La variation de la sensibilité n'est pas spécifique à un type de capteur, en effet nous observons que pour des capteurs de même type (TGS 2620, fig 2-28) cette variation de sensibilité varie d'un capteur à l'autre et d'un gaz à l'autre. En effectuant la standardisation des réponses (pour le TGS 2620, le gaz de standardisation est le NH_3), cette variation de sensibilité, éliminée pour NH_3 , pourra être réduite dans certaines limites pour les autres gaz (figure 2-29). Par contre le test de standardisation ne peut pas compenser la perte de sensibilité au méthane.

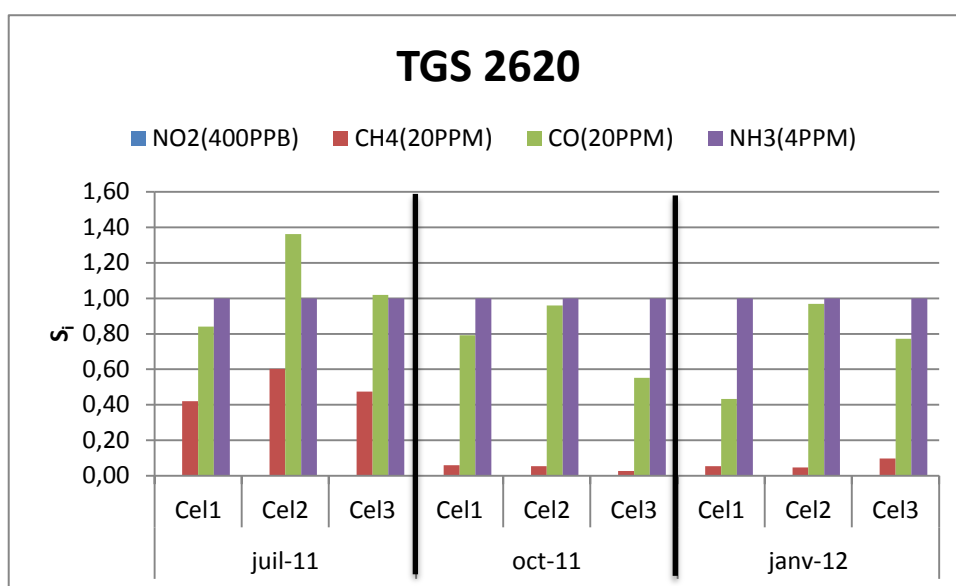


Figure 2-29 : Standardisation des réponses du capteur TGS 2620.

Les réponses du capteur TGS 2620 sont standardisées par rapport à sa réponse au NH_3 . Après standardisation, les valeurs des réponses de ce type de capteur au NH_3 sont donc ramenées à 1. Les variations de sensibilité sont ainsi atténuées pour le CO, mais non corrigées pour le méthane. Il s'agit toutefois d'un progrès important et intéressant si nous envisageons de développer des modèles robustes aux variations temporelles de sensibilité.

6. Conclusion du chapitre

Nous avons présenté dans ce chapitre les principaux éléments du système multi-capteurs conçu. Le banc de test « laboratoire » élaboré permet de caractériser parallèlement 3 cellules

de mesures, chacune contenant 9 capteurs de gaz fonctionnant en mode isotherme, à l'aide de mélanges synthétiques. Pour l'application « terrain », un banc spécifique a également été développé, tenant compte des contraintes techniques (ligne de prélèvement extérieure par pompage, autonomie de fonctionnement et stockage long terme des données).

Afin de coller au plus près aux conditions réelles de pollution atmosphérique, 7 différents gaz pour les tests en laboratoire ont été choisis dont deux d'entre eux sont des gaz oxydants, qui seront manipulés avec des basses concentrations en harmonie avec les niveaux de concentration rencontrés en air extérieur.

La dernière partie de ce chapitre a été consacrée à la mise en œuvre d'un mode opératoire pour le suivi des lignes de base et la caractérisation des performances des 9 capteurs spécifiquement choisis lors de la conception des cellules de mesure. Bien que ces capteurs présentent des avantages nombreux, nous avons vu que leur sélectivité et leur sensibilité présentent des points faibles au niveau de la dérive et de non répétabilité des sensibilités qui nécessitent des corrections. L'introduction d'un test trimestriel du suivi de la ligne de base et d'une méthode de standardisation des sensibilités contribue fortement à atténuer ces défauts. La reproductibilité inter-cellules est également améliorée par l'ensemble des outils ainsi mis en place.

Les 3 cellules ont été conçues de sorte à pouvoir intégrer l'influence de la température et de l'humidité dans des modèles, même si, dans ce travail de thèse, leur impact sera négligeable puisque les manipulations s'effectuent toujours dans des conditions environnementales contrôlées, que ce soit en laboratoire ou sur le terrain.

La capacité des 3 cellules à déterminer des signatures spécifiques d'un gaz pur, en tenant compte du niveau de concentration, y compris à faibles teneurs, a pu être établie. Cette

propriété est de bon augure pour la reconnaissance de mélanges complexes stables en milieu réel.

Dans la suite de ce manuscrit, nous présentons les résultats obtenus avec les 3 différentes cellules ainsi élaborées, et installées dans différentes stations de mesure destinées au suivi de la pollution atmosphérique. Les différents types de pollution sont analysés, des modèles de traitement des données de mesure des capteurs sont développés et les résultats obtenus sont comparés à ceux issus des réseaux de surveillance de la qualité de l'air.

Chapitre III :

Classification et indexation de la pollution de l'air avec un réseau de capteurs.

Les différents tests présentés dans le chapitre précédent ont permis de démontrer l'aptitude de notre système à détecter différents polluants gazeux dans l'air avec des critères métrologiques répondant aux besoins des applications terrains envisagées (mesure des faibles teneurs...). Ces systèmes ont été installés dans différentes stations de mesure de la qualité de l'air afin d'établir s'il était possible de mettre en évidence une signature décrivant la pollution atmosphérique spécifique de la zone concernée et de la période étudiée. L'identification de ces signatures doit être obtenue en corrélant les réponses des capteurs aux mesures des analyseurs installés dans ces stations.

Dans la première partie de ce chapitre, sont décrits les différents sites de mesures étudiés. Nous chercherons dans un 1^{er} temps à caractériser les différents types de pollution rencontrés sur ces sites, ensuite nous décrirons l'analyse utilisée pour définir et visualiser ces pollutions, et les méthodes utilisées pour en extraire des informations pertinentes. Enfin, nous

présenterons les différentes analyses mises en œuvre pour obtenir des modèles de comportements capables de prédire des informations sur les pollutions à partir des réponses des capteurs.

1. Tests terrain

Après avoir validé notre système en laboratoire, l'étape suivante consiste à réaliser des mesures sur des sites pollués. Nous avons décidé d'étudier la pollution durant deux saisons différentes, une en été pour pouvoir être capable de rencontrer des épisodes de la pollution photochimique et l'autre en automne. Pour chaque campagne nous avons choisi d'installer nos cellules contenant nos systèmes multi-capteurs dans plusieurs stations de mesure représentatives de différents types de sites de surveillance de la pollution atmosphérique.

Campagne d'été :

Pour cette campagne nous nous sommes installés dans le département de l'Oise situé en région Picardie. Etant donnée sa position au nord de la région Ile de France, ce département est soumis à de fréquents épisodes de pollution photochimique. Cette campagne, qui a duré 4 semaines du 27 juillet 2011 jusqu'au 24 août 2011, a été effectuée en collaboration avec ATMO-Picardie, l'association de surveillance de la qualité de l'air en Picardie. Deux sites ont été choisis comme sites de mesures : les villes de Creil et de Nogent sur Oise, le premier étant un site urbain et l'autre un site périurbain.

Campagne d'automne :

Pour cette campagne, les réseaux des capteurs ont été installés en région du Nord - Pas de Calais. Les mesures ont été effectuées en collaboration avec ATMO-Nd-PdC, l'association de surveillance de la qualité de l'air dans la région Nord – Pas de Calais. 5 sites dont 2 urbains, 2 trafics, et 1 périurbain, ont été choisis entre le 3 octobre 2011 et le 14 décembre 2011. Ils sont situés respectivement à Douai, Béthune, Roubaix, Valenciennes, et Saint Amand les Eaux.

Le tableau 3-1 récapitule les propriétés des différentes stations de l'ensemble des sites choisis. Par ailleurs, la figure 3-1 présente la localisation géographique de ces stations de mesures.

Tableau 3-1 : Caractéristiques des stations de mesures choisies pour les campagnes de terrain.

Période	Site	Type de station	Polluants mesurés par ATMO
Eté	Creil	Urbain	NO-NO ₂ -O ₃ -PM _{2,5} -PM ₁₀ -CO-SO ₂
	Nogent	Périurbain	NO-NO ₂ -O ₃ -PM ₁₀
Automne	Roubaix	Trafic	NO-NO ₂ -CO-PM ₁₀
	Valenciennes	Trafic	NO-NO ₂ -PM ₁₀
	Saint Amand	Périurbain	NO-NO ₂ -O ₃
	Douai	Urbain	NO-NO ₂ -O ₃ -PM _{2,5} -PM ₁₀
	Béthune	Urbain	NO-NO ₂ -O ₃ -PM _{2,5} -PM ₁₀

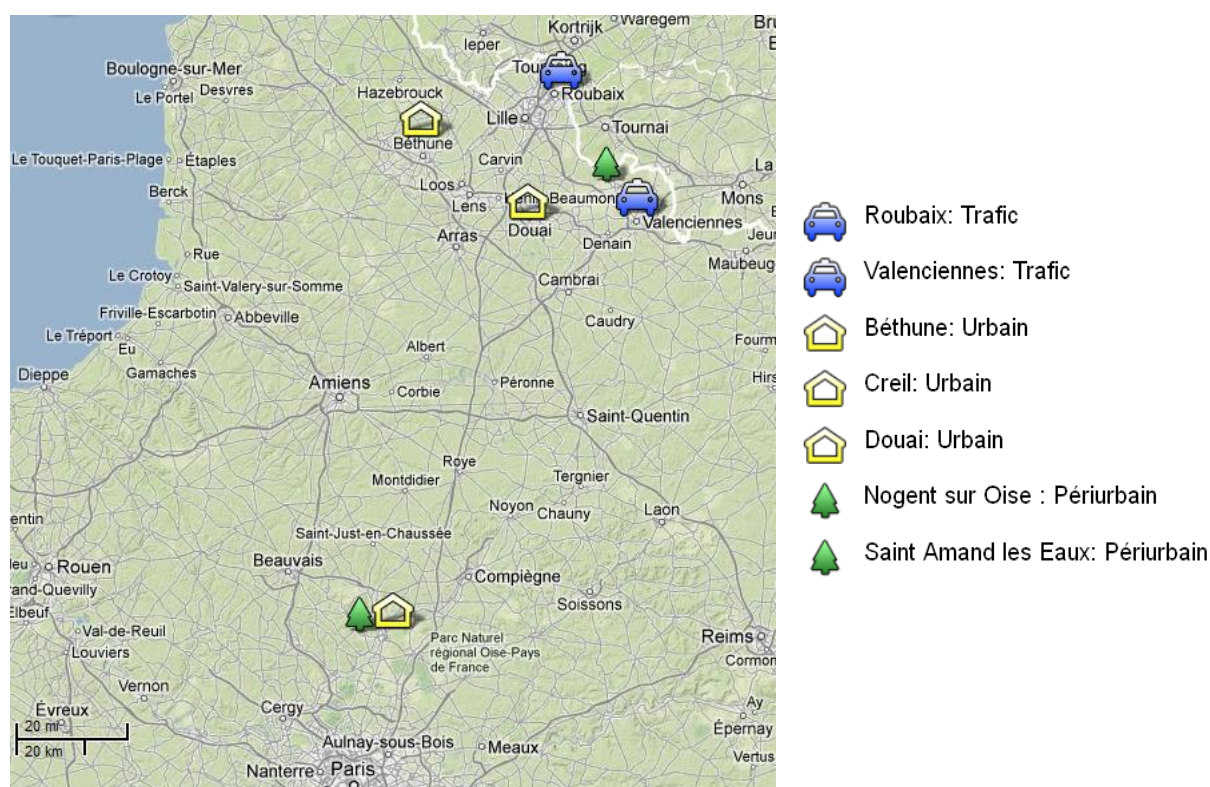


Figure 3-1 : Plan détaillant la position des différents sites de mesures.

La durée de chaque campagne de mesure dans chaque station est de 7 jours, afin d'enregistrer les différents événements hebdomadaires susceptibles de se produire. Le temps de réponse des capteurs de gaz choisis est toujours inférieur à 90 secondes, c'est pourquoi les points de mesures ont donc été enregistrés toutes les 2 minutes.

2. Caractéristiques des pollutions atmosphériques des sites étudiés

Pour effectuer une corrélation pertinente entre les réponses des capteurs et les données physico-chimiques des polluants atmosphériques issues des stations de mesures, une étude appropriée des différents événements et caractéristiques de la pollution atmosphérique sur les sites de mesure choisis est devenue indispensable. Dans ce but, nous avons travaillé sur un historique d'une année (2010) de données des concentrations des polluants dans l'atmosphère issue de chaque station de mesure étudiée. Les gaz retenus pour cette étude sont le NO, NO₂ et O₃, car leurs données sont complètes sur toutes les stations, cependant, pour les stations trafic, la mesure d'O₃ n'est pas pertinente (cf. Chapitre 1 partie 2.3).

2.1. Analyse des concentrations des polluants par type de station

L'analyse de cet historique a été orientée pour trouver des formes répétables et des caractéristiques qui peuvent définir et distinguer un type de pollution d'un autre. Nous avons étudié les formes journalières, hebdomadaires et saisonnières de ces données. Une étude statistique nous a permis de montrer le caractère répétable des données inter-stations. Les résultats sont présentés dans le tableau 3-2.

Tableau 3-2 : Etude statistique effectuée sur les différentes stations de mesure : moyenne, médiane, maximale et écart-type.

Station		NO (µg/m ³)				NO ₂ (µg/m ³)				O ₃ (µg/m ³)			
Station	Type	Moy.	Med.	Max	σ	Moy.	Med	Max	σ	Moy.	Med	Max	σ
Creil	Urbaine	9	3	243	16	24	21	122	17	42	40	235	29
Douai	Urbaine	9	2	296	18	25	23	100	15	39	38	182	27
Béthune	Urbaine	6	2	191	13	22	18	150	16	44	43	182	27
Nogent	Périurbaine	12	5	254	21	26	22	139	17	43	40	236	30
Saint Amand	Périurbaine	6	1	230	16	14	12	69	10	37	39	124	23
Roubaix	Trafic	32	16	801	46	46	42	301	25	---	---	---	---
Valenciennes	Trafic	20	10	405	29	37	34	152	18	---	---	---	---

Le tableau 3-2 présente les propriétés statistiques des polluants NO, NO₂ et O₃ durant une période d'un an. Nous remarquons que les sites urbains et périurbains étudiés présentent

des valeurs statistiques similaires en O_3 ce qui est logique compte tenu du caractère régional des épisodes de pollution photochimique à l'origine des concentrations en O_3 . Les concentrations en NO et NO_2 observées dans les sites trafics sont plus importantes que celles des sites urbains et périurbains étudiés. Rappelons que l' O_3 est quasiment absent des sites « trafic » et n'est d'ailleurs pas mesuré à cause des fortes concentrations de NO dans l'air qui vont potentiellement détruire l' O_3 présent. Nous pouvons observer cette similitude de comportement entre les différents types de station via les graphes des moyennes annuelles des polluants présentés en la figure 3-2. Nous montrons par ailleurs sur la figure 3-3 la corrélation entre les concentrations de NO et de NO_2 dans un site trafic (Roubaix).

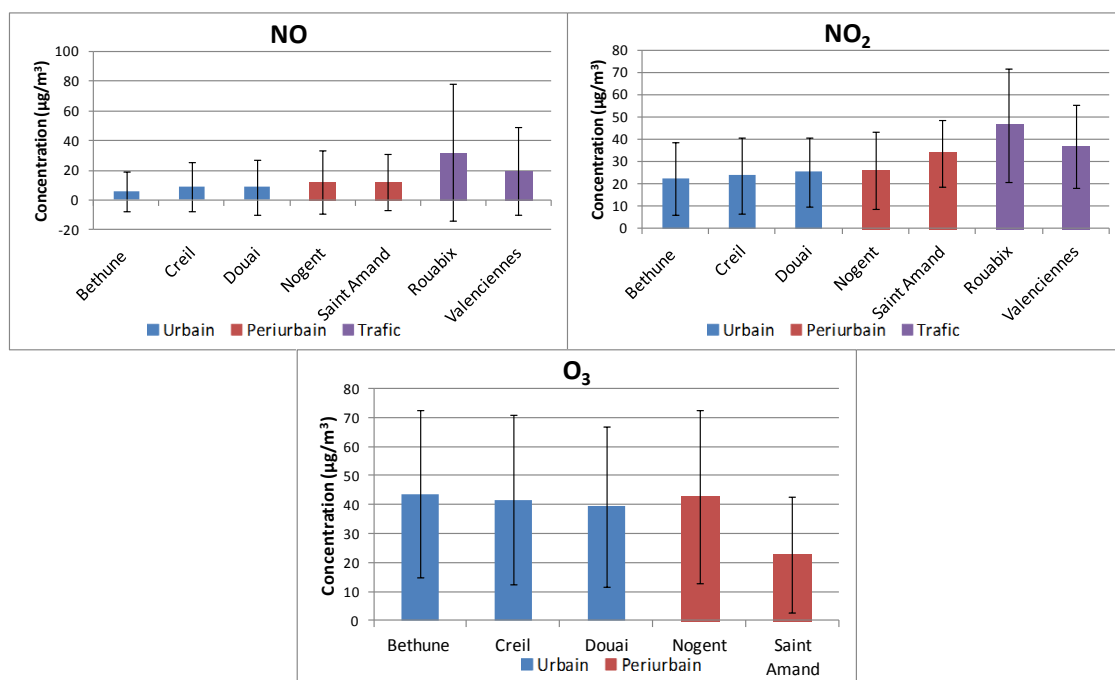


Figure 3-2 : Les moyennes et les écarts types annuels des concentrations en NO, NO₂ et O₃ dans les différentes stations [2010].

Afin de différencier les sites et d'identifier leurs caractéristiques, les figures 3-3 et 3-4 présentent les corrélations entre les concentrations de NO et NO₂ sur différents sites d'études.

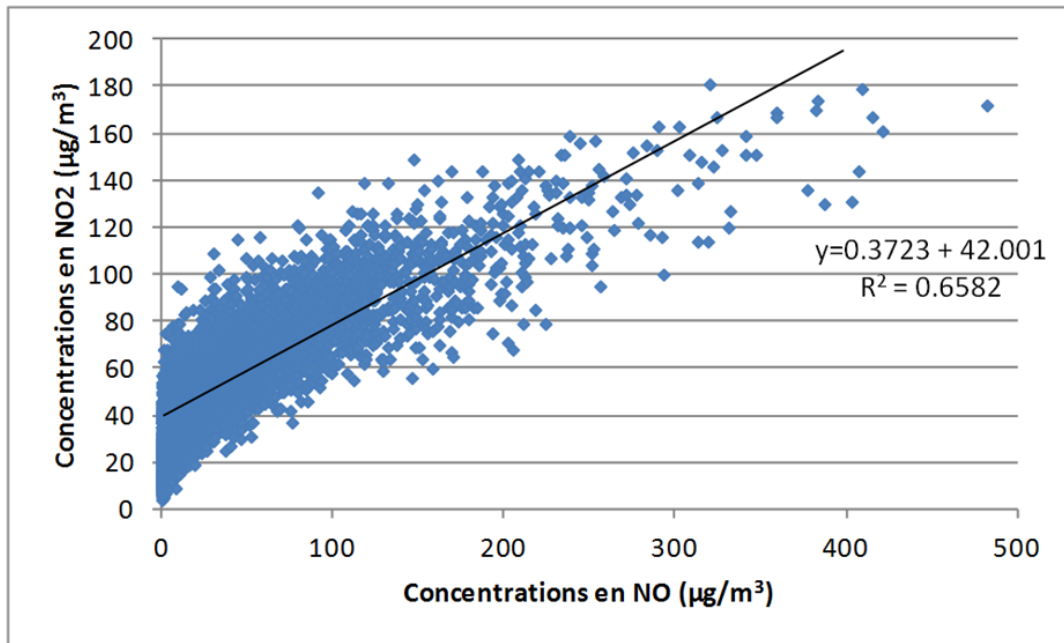


Figure 3-3 : Corrélation entre les concentrations de NO et de NO₂ dans un site trafic (Roubaix) durant un an, y et R² établies pour NO > 20 µg/m³.

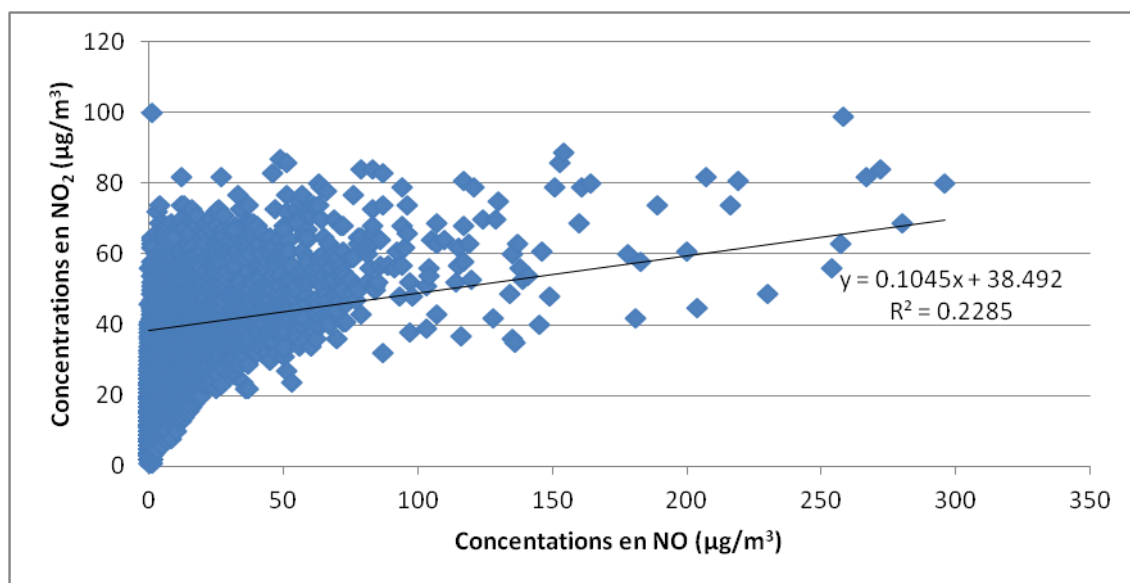


Figure 3-4 : Corrélation entre les concentrations de NO et de NO₂ sur un site urbain (Douai) durant un an, y et R² établies pour NO > 20 µg/m³.

Ainsi, pour ce qui concerne les sites trafic, nous pouvons observer qu'une forte concentration en NO est toujours accompagnée d'une haute concentration en NO₂, constat généralement observé aux heures de pointes de trafic. Ainsi il apparaît que lorsque la concentration en NO dépasse 20 µg/m³ la relation entre NO et NO₂ présente une bonne

corrélation (par exemple $R^2=0.65$ pour le site de Roubaix). Par contre cette relation est bien moins nette sur un site urbain ou périurbain comme montré dans la figure 3-4: la présence d'une forte concentration en NO_2 est indépendante de la présence de NO dans l'atmosphère. Les sources de trafic sont moins importantes dans ces zones, c'est pourquoi les concentrations en NO sont relativement basses durant l'année. Le tableau 3-2 rend bien compte de cet état de fait puisque les moyennes de concentration en NO et NO_2 dans les sites urbains et périurbains sont inférieures à celles observées sur les sites trafics.

La figure 3-5 montre la relation inversement proportionnelle entre NO et O_3 sur le site urbain de Douai : une forte concentration en O_3 est toujours accompagnée d'une faible concentration en NO, et inversement.

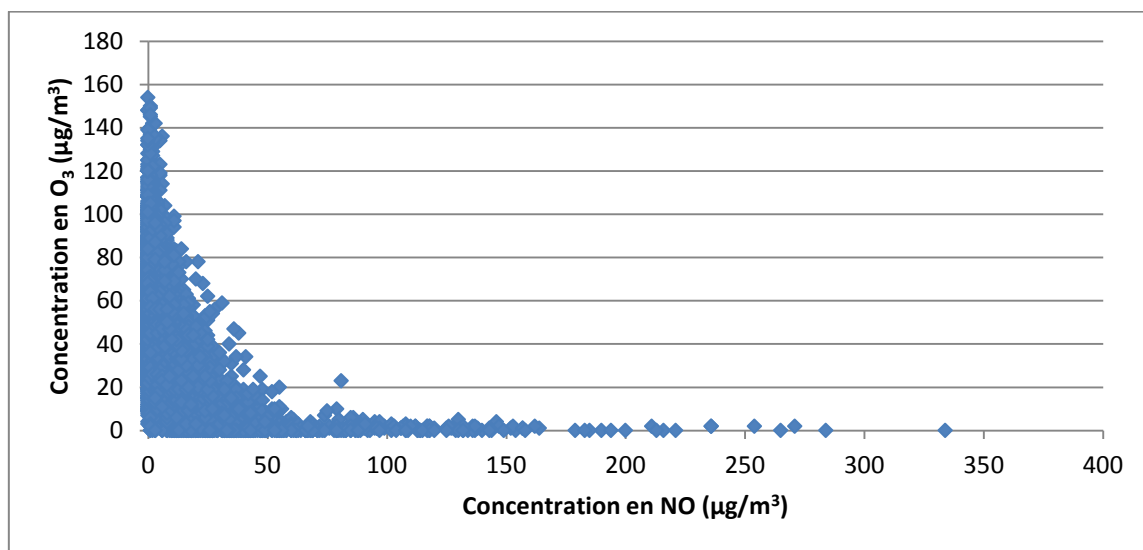


Figure 3-5 : Corrélation entre les concentrations d' O_3 et de NO sur le site de Douai (Urbain).

Les relations montrées dans les figures 3-3, 3-4, et 3-5 sont observées sur toutes les stations. NO et NO_2 sont fortement corrélés dans les sites trafics, et O_3 et NO sont anti-corrélés dans les sites urbains et périurbains. Les figures présentant les relations entre les polluants dans les différentes stations de mesures sont présentées dans l'annexe B page 196.

2.2. Identification de signatures

Nous nous intéressons ici aux différentes allures de présence des polluants en fonction de périodes journalières, hebdomadaires, et saisonnières. La présence de formes répétables enrichit nos informations sur les pollutions au sens où elles peuvent être utilisées pour identifier la principale source de la pollution rencontrée sur un site d'étude. Ces informations peuvent également être utilisées comme variables additionnelles dans les modèles de prédiction qui seront développés par la suite. Afin d'étudier l'influence éventuelle de la saisonnalité sur les allures journalières et hebdomadaires, nous avons séparés les données en 2 groupes : un groupe « estival », constitué des mesures enregistrées entre le 1^{er} avril et le 31 octobre, et un groupe « hivernal » entre le 1^{er} novembre et le 31 mars.

2.2.1. Les formes journalières

En étudiant la forme journalière des données issues des analyseurs des stations de surveillances, nous observons des aspects répétables entre les mesures des polluants provenant des sites urbain et périurbain d'une part et des sites trafic d'autre part, deux formes sont remarquables durant les deux saisons différentes.

Prenons comme exemple le site urbain de Creil. La figure 3-6 représente la forme journalière des concentrations de NO, NO₂ et O₃ rencontrées sur ce site durant la période hivernale. Les concentrations de NO et NO₂ sont maximales vers 8h le matin et vers 19h le soir ce qui correspond aux horaires des pics de trafic quotidiens. D'autre part, la forme journalière d'O₃ confirme une anti-corrélation avec celle de NO et NO₂.

La concentration en O₃ est minimale durant les heures de pointe en trafic correspondant aux maximums journaliers en NO et NO₂, et elle est maximale vers 14h où le minimum de concentration de NO et NO₂ est enregistré. A noter également que la présence d'O₃ est

corrélée aux horaires de forts ensoleillements : cette concentration est maximale en pleine journée et minimale durant la nuit.

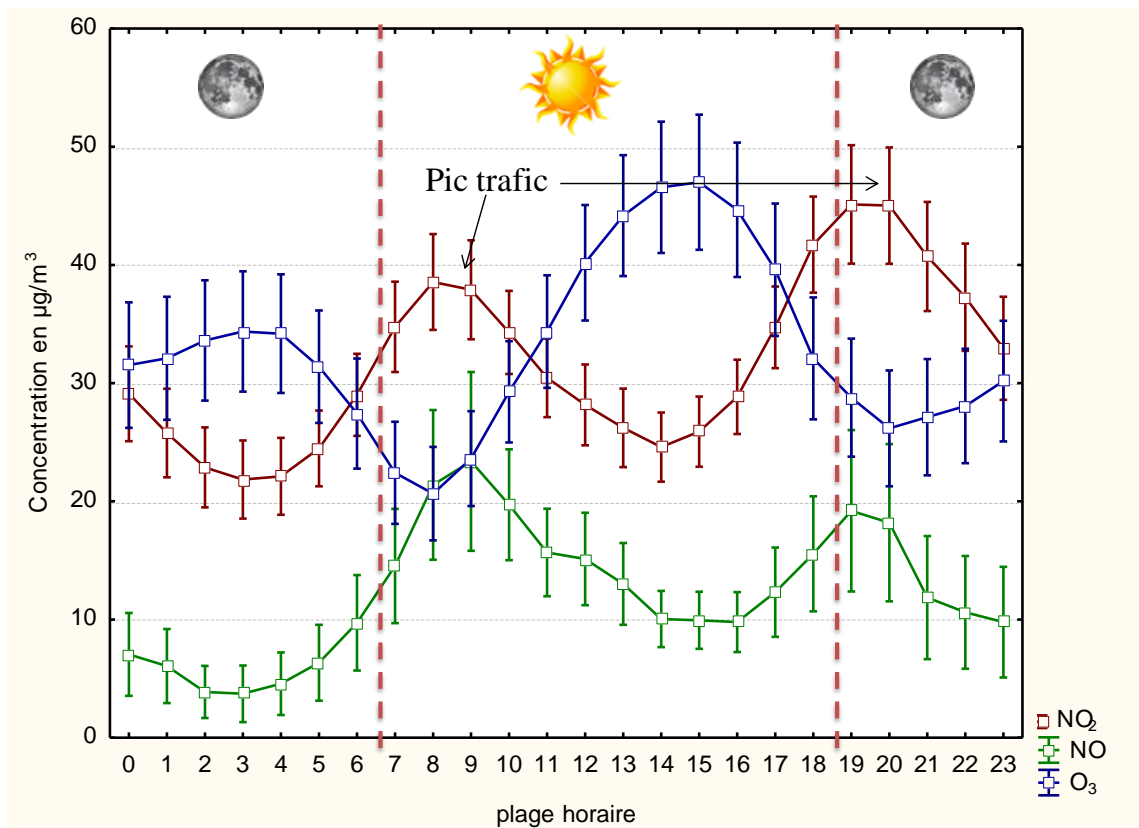


Figure 3-6 : Forme journalière durant la saison hivernale de la présence de NO et NO₂ dans le site urbain de Creil.

En comparant les allures diurnes des concentrations en polluants en période hivernale à leurs allures en période estivale, il est aussi possible d'identifier une forte différence (voir figure 3-6 et 3-7). Dans un premier temps, les concentrations en oxydes d'azotes (NO et NO₂) sont beaucoup plus fortes en saison hivernale tandis que les concentrations en O₃ sont sensiblement plus fortes en saison estivale. Trois éléments pouvant expliquer ces différences :

- D'une part, les phénomènes de dilution sont beaucoup plus importants en période estivale (hauteur de la couche de mélange plus importante) et donc si on considère les sources à l'origine des concentrations d'intensité comparables cela aura pour conséquence des concentrations ambiantes plus faibles.

- D'autre part, en période hivernale, il peut y avoir une augmentation de l'intensité des sources à l'origine des concentrations en NO et NO₂ comme par exemple les phénomènes de combustion domestiques liées au chauffage.
- De plus, l'activité photochimique est plus intense entraînant une production des polluants secondaires tels que l'O₃ bien plus importante en particulier en milieu d'après-midi.

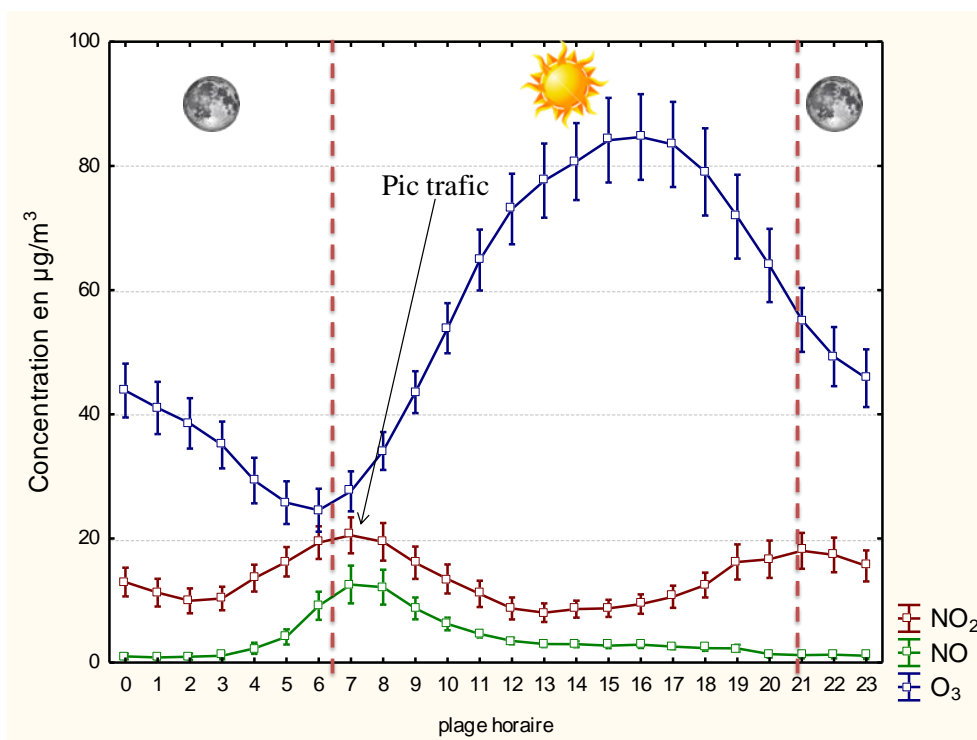


Figure 3-7 : Forme journalière des concentrations de NO, NO₂ et O₃ de la saison estivale sur le site urbain de Creil.

Dans un site trafic, et comme montré dans la partie 2.3 du chapitre 1, il n'est pas utile de suivre l'évolution de l'O₃ dans l'atmosphère à cause des fortes émissions en NO et NO₂ issus essentiellement des échappements des véhicules, ce qui entraîne la titration de l'O₃ par NO et a pour conséquence la production en particulier de NO₂ et de COV oxygénés (COV-O). Nous montrons dans la figure 3-8 l'évolution de la forme journalière en période hivernale des concentrations en NO et NO₂ sur le site trafic de Roubaix.

Les pics de trafic observés correspondent aux mêmes horaires de pointe rencontrés dans les sites urbains et périurbains, par contre les niveaux de concentration sont plus importants

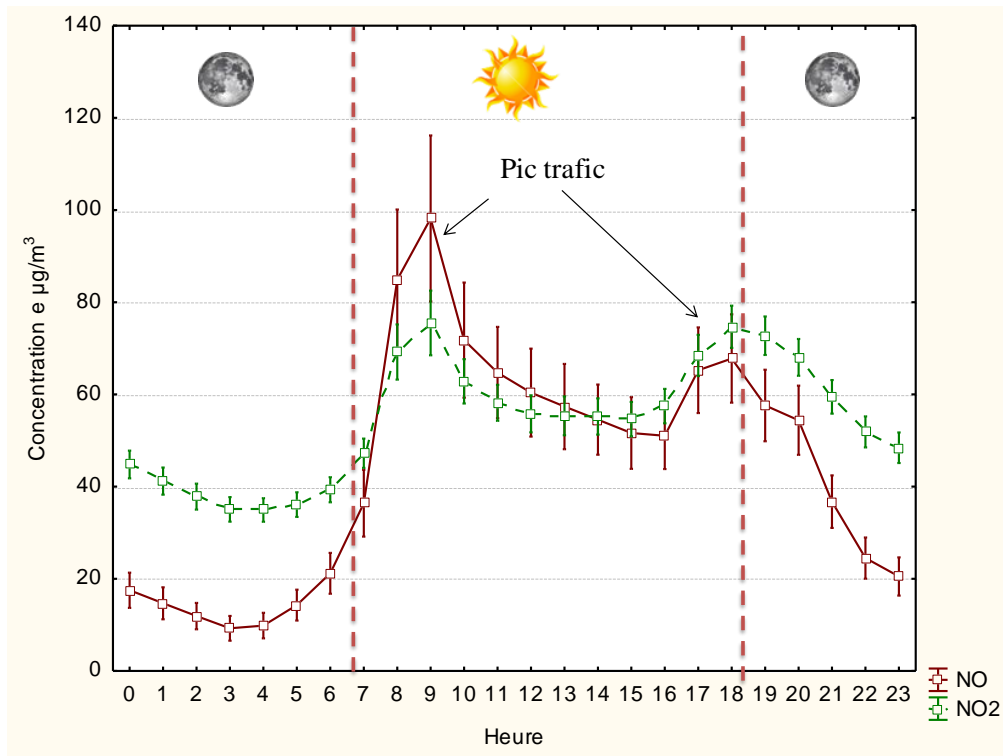


Figure 3-8 : Evolution journalière des concentrations en NO et NO₂ de la saison hivernale sur le site trafic de Roubaix.

La similarité des formes journalières de NO et NO₂ dans les sites trafics indique qu'ils proviennent de la même source ou au moins du même type de source. Cette similarité peut ainsi être considérée comme une caractéristique des sources trafic.

Nous pouvons donc corréler la présence d'une forte concentration simultanée en NO et NO₂ à une source trafic ainsi qu'une forte concentration d'O₃ à une source photochimique. Dans les sites urbains et périurbains, l'absence d'une source trafic importante implique une concentration faible en NO voire l'absence de ce polluant dans certains cas. On se retrouve alors soit avec une pollution photochimique en cas d'excès d'O₃ soit avec une forte concentration en NO₂ en cas d'absence d'O₃. Comme nous avons défini une pollution photochimique par la présence d'O₃, nous pouvons ainsi définir une pollution urbaine comme la présence d'une forte concentration de NO₂ non accompagnée par NO. **Donc l'étude de ces**

trois polluants NO, NO₂ et O₃ va nous permettre de définir différents types de pollution atmosphérique. Nous montrons dans l'annexe C page 203 les formes journalières observées les différents sites de mesure.

2.2.2. Les formes hebdomadaires :

Afin d'étudier l'effet du jour de la semaine sur la pollution, nous avons à nouveau séparé les données entre la saison hivernale et la saison estivale. En étudiant l'évolution des concentrations de NO, NO₂ et O₃, nous avons identifié un lien entre les moyennes journalières des polluants et le jour de la semaine. Durant un jour ouvré (jour non férié et hors week-end), la concentration en NO et NO₂ est plus élevée que lors des jours non ouvrés. Ce phénomène est inversé pour l'O₃. Durant les jours caractérisés par une circulation trafic réduite, comme les jours non ouvrés et les week-ends, les émissions de NO, CO, et COV provenant des sources de circulations trafic diminuent, en revanche l'O₃ augmente. Pour cela, nous remarquons que les concentrations en O₃ sont plus élevées durant les jours non ouvrés que les jours ouvrés, en raison de la baisse de la circulation automobile et d'autres activités anthropiques. La figure 3-9 illustre le cas de la station urbaine de Creil, les évolutions des moyennes journalières des concentrations de NO, NO₂ et O₃ en saison hivernale y sont reportées.

Durant une période estivale, nous observons que cette différence entre les jours ouvrés et les jours non ouvrés est réduite. Les concentrations des oxydes d'azotes (NO et NO₂) en période estivale sont moins fortes qu'en période hivernale, contrairement aux concentrations d'O₃ qui sont plus fortes en période estivale. La figure 3-10 montre l'évolution des concentrations des polluants sur le site urbain de Creil durant la période estivale durant l'année 2010.

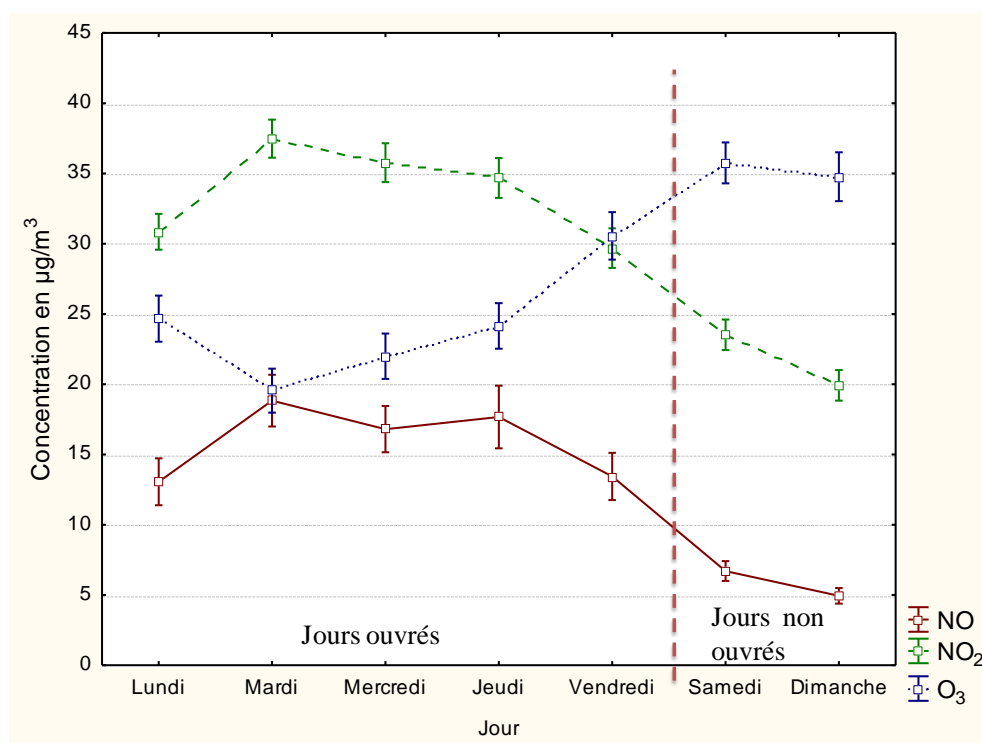


Figure 3-9 : Tracé des moyennes d'O₃, NO et NO₂ hebdomadaire en période hivernale sur le site de Creil.

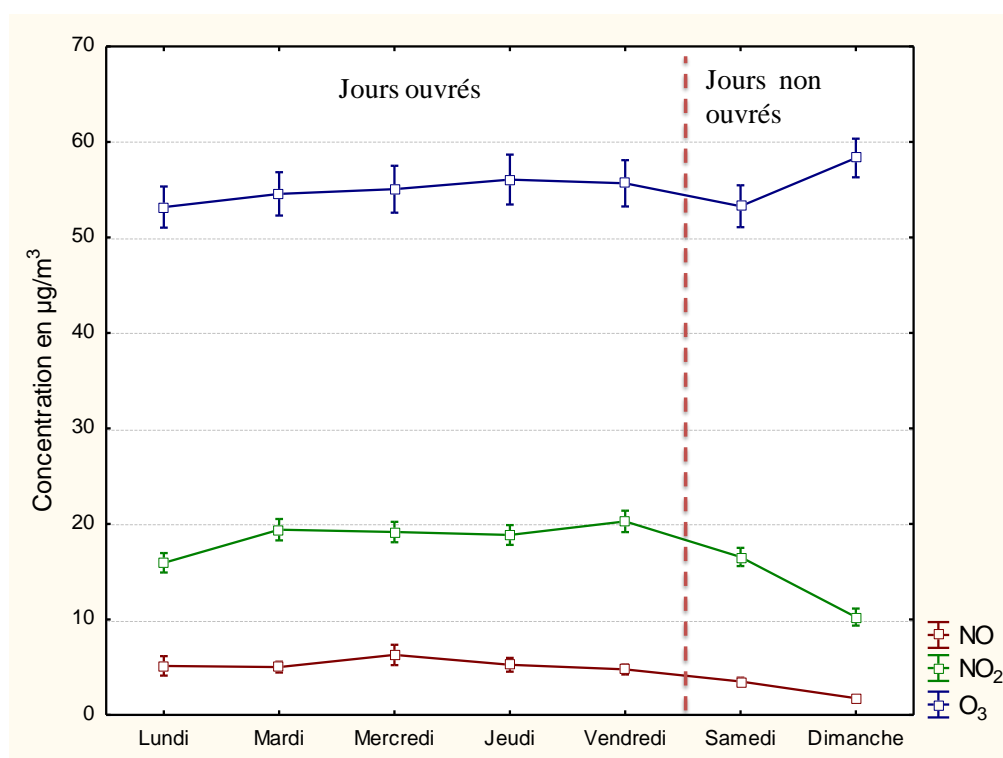


Figure 3-10 : Tracé des moyennes d'O₃, NO et NO₂ par jour en période estivale sur le site de Creil.

La notion de jours « ouvrés » et jours « non ouvrés » constitue donc une information additionnelle pour définir les types de pollutions, à utiliser lors de l'établissement des modèles

de prédiction en plus de l'information horaire, et en tenant compte de l'effet de saisonnalité. Nous montrons dans l'annexe C page 203 les formes hebdomadaires observées sur les différents sites de mesures étudiés qui confirment que les comportements des différents polluants que nous présentons ici en exemple peuvent être généralisés.

3. Démarches de modélisation-prédiction

Dans ce qui suit, nous présentons en détail les méthodes de modélisation mises en œuvre pour prédire des informations sur la pollution atmosphérique à partir des données des capteurs, que ce soit de façon catégorielle (classification) ou quantitative. Nous avons vu dans le chapitre 1 les différentes méthodes multi-variables de modélisation qui existent. La plupart de ces méthodes (ex : PCA, PCR, PLS) sont des méthodes linéaires et peu adaptées à la réponse non linéaire des capteurs. Nous sommes ainsi amenés à utiliser préférentiellement les réseaux de neurones artificiels (ANN) qui permettent de prendre en compte la non-linéarité des capteurs. Dans ce but, nous utilisons le logiciel, STATISTICA qui nous donne la liberté de choisir l'architecture des réseaux avec une gamme de choix des fonctions d'activation (voir chapitre1, réseaux de neurones).

3.1. Prétraitements de données

Dans le but d'optimiser les modèles de prédiction, nous introduisons une nouvelle normalisation des valeurs inter-capteurs basée sur la normalisation introduite par A.-C Romain (Romain, et al., 2005). Cette nouvelle normalisation, appliquée aux différents capteurs d'une même cellule, est la suivante :

$$Z_i = \frac{S_i}{\sqrt{\sum_{j=1}^k S_j^2}} \quad (1)$$

Avec :

- Z_i la valeur de la conductance dite normalisée du capteur i
- S_i et S_j la valeur de la conductance standardisée en sensibilité du capteur i définie précédemment par $S_i = \frac{X_i}{X_{iC}}$

Le but d'une telle normalisation est l'atténuation des informations redondantes dans les réponses des capteurs au profit de l'amplification des écarts entre eux. Ce traitement n'a de sens que si les réponses des différents capteurs d'une même cellule sont du même ordre de grandeur, c'est la raison pour laquelle nous appliquons ce traitement à la variation de conductance standardisée en sensibilité S (Voir chapitre 2, partie 5.3).

Il est utile de rappeler ici les différents traitements successifs appliqués aux mesures de conductance : Le schéma de la figure 3-11 rappelle le bilan des traitements subis par les réponses des capteurs. Les réponses brutes sont obtenues en mesurant les conductances des couches sensibles de chaque capteur. D'une part, l'étude de la variation relative de conductance de chaque capteur sous un air de référence, permet de s'affranchir des variations des lignes de base des capteurs. D'autre part, sous un mélange de gaz polluants de concentrations stables, les réponses standardisées sont calculées pour homogénéiser et rapprocher les sensibilités des capteurs. Finalement, les réponses des capteurs normalisées sont calculées à partir des réponses standardisées. Ces trois traitements (« variation relative », « standardisation » puis « normalisation ») seront systématiquement utilisés, sauf mention contraire, pour la suite de nos modélisations. Nous consacrons la partie 5.1.1. pour étudier l'effet de ces traitements sur la modélisation.

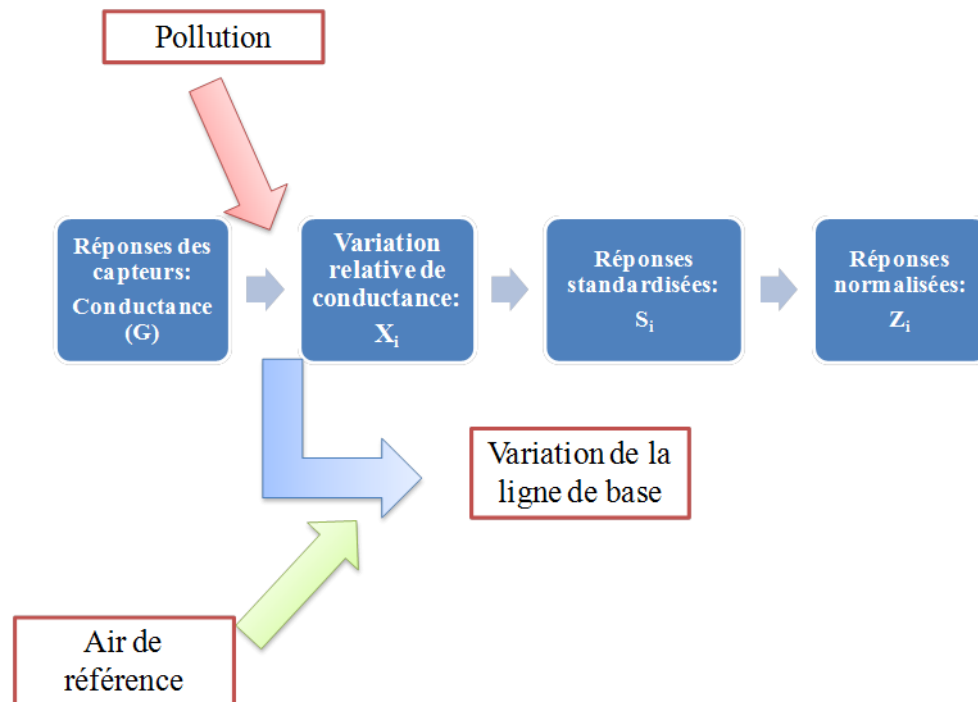


Figure 3-11 : Bilan de tous les traitements des réponses des capteurs.

Durant les mesures effectuées par notre système, l'acquisition complète des données des capteurs se fait toutes les 2 minutes. Pour une semaine de mesure nous obtiendrons donc un nombre approchant les 5040 observations. Par contre, les données physico-chimiques des concentrations des polluants réglementés dans l'air ont été fournies par les associations de surveillance en une périodicité horaire soit 168 observations par semaine. Nous avons décidé de calculer des moyennes horaires pour les réponses de chaque capteur pour obtenir des fichiers de données synchronisés.

3.2. Elimination des points aberrants

Lorsque nous observons les données capteurs avant moyennage, nous pouvons observer quelques points aberrants. Ces points aberrants peuvent provenir d'un évènement spécifique aléatoire comme le passage d'un véhicule à proximité d'une station ou l'émission ponctuelle d'odeur n'appartenant pas à la pollution atmosphérique. Ceci induit des pics brutaux dans les réponses des capteurs comme montré dans la figure 3-12. Ces points seront

enlevés du jeu de données avant que ces données ne soient moyennées et synchronisées avec les données physicochimiques.

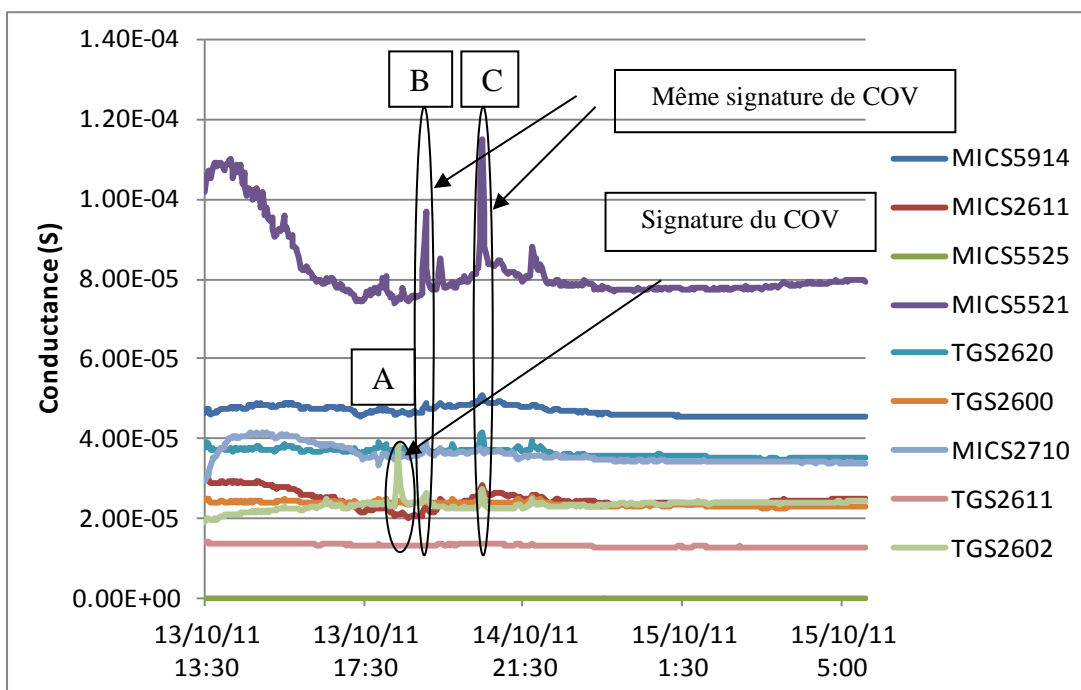


Figure 3-12 : Exemples de valeurs anormales enregistrées par les capteurs

Cette figure représente la conductance des capteurs durant la campagne sur le site de Saint-Amand. Une déchetterie est située à proximité de cette station. Les pics B et C enregistrés par les capteurs ont probablement pour origine la même source puisque les réponses des capteurs de COV présentent la même évolution relative. Il s'agit sûrement d'une source d'hydrocarbure. Le pic A est caractérisé par une réponse forte unique du capteur TGS2602, absente des autres capteurs, donc ce phénomène peut indiquer l'existence de toluène ou d'une source odorante. Il est possible de reconnaître la source de ces pics en connaissant la signature de la pollution sur les capteurs, dans ce cas le point A peut avoir comme origine l'évacuation de pot de peinture à l'arrière de la station, par contre une telle reconnaissance nécessite une étude préalable et complexe dans le laboratoire, focalisée sur les différentes signatures de polluants et des odeurs. Etant donné la complexité d'une telle étude car ces pics observés sur nos capteurs n'appartiennent pas à l'étude de la pollution atmosphérique que l'on souhaite caractériser, nous avons décidé de considérer ces pics

comme points aberrants et de les éliminer de nos modèles. Nous verrons en fin de chapitre qu'il peut être possible de détecter de tels évènements en phase de prédiction.

3.3. Choix des variables

Les modèles de comportement que nous essayons d'établir permettent de modéliser des sorties (Y variables) en fonction d'entrées (X variables). Dans notre cas, nous cherchons à prédire les informations sur la pollution à partir des réponses des capteurs, et éventuellement d'autres variables facilement disponibles.

Dans la partie consacrée à l'étude des données de pollution atmosphérique nous avons trouvé des informations additionnelles qui peuvent aider à identifier des types de pollution. Des variables supplémentaires ont donc été ajoutées au jeu de signaux capteurs dans le but d'incorporer le plus d'informations possible :

- la réponse du capteur d'humidité, pour intégrer l'effet de l'humidité dans les modèles,
- une variable décrivant l'heure du jour, c'est une variable qui varie entre 0 et 12 (croissante de 0h à 12h puis décroissante jusqu'à minuit comme montre la figure 3-13) pour incorporer les formes journalières de pollution.
- une variable discrète indiquant les jours ouvrés et les jours non ouvrés. C'est une variable ayant une valeur 0 pour les jours ouvrés, et une valeur 1 pour les jours fériés, les samedis et les dimanches. Le but de cette variable est d'ajouter une information sur la forme hebdomadaire d'une pollution.

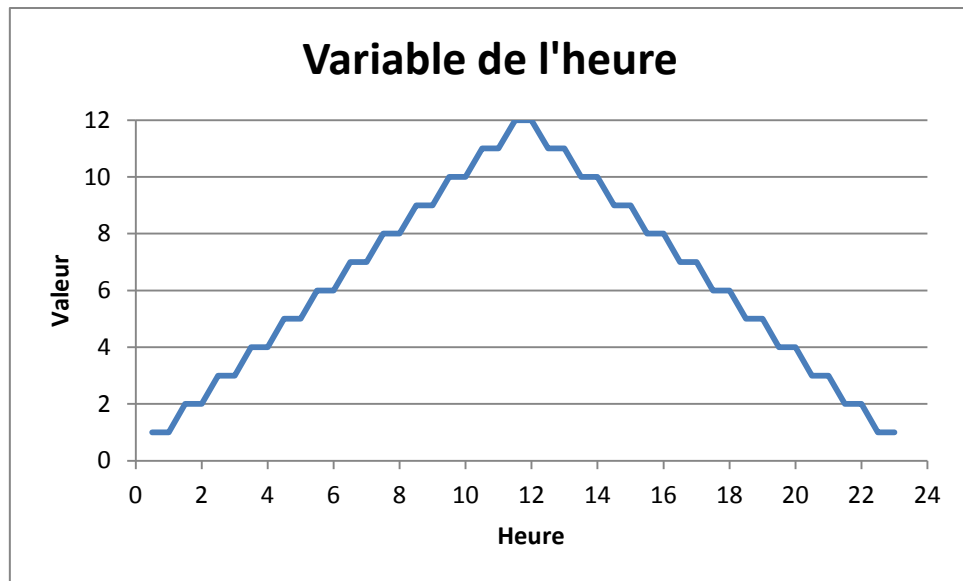


Figure 3-13 : Variable de l'heure choisie à incorporer dans les modèles à calculer.

Nous avons vérifié que ces 3 variables sont bien sûr insuffisantes en elles mêmes pour caractériser une signature de pollution, mais améliorent quelque peu les performances de prédiction. Une étude, exposée plus loin, montre l'effet de ces variables supplémentaires sur la performance des modèles (§5.1.2.). La température de l'air n'a pas été incorporée puisque les stations de mesure sont climatisées et maintenues à une température constante de 23°C et que les réseaux sont placés dans ces stations.

3.4. Apprentissage des modèles, limites de validité

Dans un premier temps, nous avons débuté en cherchant un modèle calculé à partir d'une seule cellule pour étudier le caractère modélisable du système, puis nous poursuivons le développement de modèles inter cellules (2 puis 3 cellules). L'architecture des modèles à base des réseaux de neurones est présentée sur la figure 3-14. Elle est composée de 12 entrées contenant les X variables qui sont : les réponses des 9 capteurs de gaz, l'humidité de l'air, la variable sur les jours ouvrés, et la variable horaire. En sortie nous disposons Y variables, qui peuvent être les concentrations des polluants ou les classes de pollution.

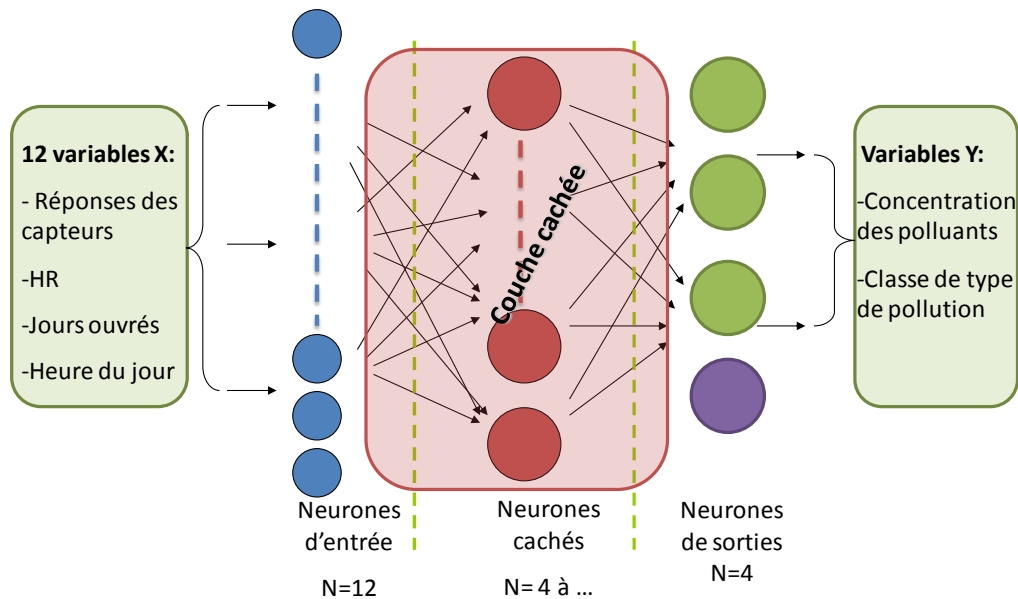


Figure 3-14 : Architecture générale des réseaux de neurones à calculer.

L'utilisation des réseaux de neurones nécessite 3 jeux de données comme expliqué dans le chapitre I-4.3.3 :

- un jeu d'apprentissage pour l'établissement du modèle
- un jeu de validation pour valider et éviter une sur modélisation
- un jeu de test pour évaluer les performances d'un modèle.

Le premier jeu sera utilisé pour construire le modèle et le faire converger grâce à l'algorithme d'optimisation, cette convergence sera vérifiée lors des itérations sur le jeu de validation, et l'apprentissage interrompu dès que les performances sur ce dernier jeu ne progresseront plus (RMSEP arrête de diminuer). Le modèle terminé sera enfin testé avec le jeu de test afin d'en évaluer les performances. Si une campagne de mesure a une durée de 7 jours, nous avons décidé de diviser nos jeux de données de cette façon : 5 jours pour l'apprentissage, 1 jour pour la validation et 1 jour pour le test. Ainsi, 60% des données seront utilisées pour l'apprentissage, 20% pour la validation et 20% pour le test.

Notons ici que la plupart des logiciels d'ANN, comme Statistica, permettent un choix automatique et aléatoire de ces 3 jeux de données. Cette méthode n'est pas valide dans notre

cas car il faudrait que toutes les données soient indépendantes dans le temps comme montré dans le chapitre 1 paragraphe 4.4.3. Or nos données sont horaires, et les phénomènes de pollution atmosphérique ont des constantes de temps qui le plus souvent sont de plusieurs heures. Donc nous imposons un ordre chronologique dans les 3 jeux de données. Ainsi le groupe de test sera composé des données sur 24 heures successives, également, le groupe de validation sera aussi composé des données sur d'autre 24 heures successives dans le but d'assurer une indépendance temporelle entre les jeux de données de différents groupes.

Rappelons que les ANN peuvent être utilisés soit en quantification soit en classification (ou catégorisation) :

- En quantification, les cibles d'apprentissage (et donc les sorties) sont des variables pouvant évoluer continument, comme par exemple des concentrations.
- En classification, les cibles d'apprentissage sont des variables booléennes de valeur zéro ou un. Lors de la prédiction, les sorties ont des valeurs variant continument, la classe choisie est alors celle présentant le poids le plus fort.

Nous utilisons le module SANN du logiciel Statistica pour ces modélisations. Ce module calcule un nombre de réseaux choisi automatiquement ou prédéfini par l'utilisateur (notre choix varie entre 100 et 500). Le module teste ainsi aléatoirement un certain nombre de paramètres comme le nombre de neurones cachés du modèle et les fonctions de transfert utilisées par ces neurones. Les réseaux retenus par le programme sont fonction de la performance d'apprentissage, validation et test. Deux paramètres peuvent être utilisés pour évaluer les performances d'un modèle :

- Le RMSEP (*Root Mean Squared Error of Prediction*) qui quantifie l'écart entre les cibles et les sorties du réseau, utilisable principalement en quantification, mais aussi en classification. $\left(RMSEP = \sqrt{\frac{\sum (Y - Y')^2}{n}} \right)$
- Taux de classification réussie, en classification uniquement, c'est le pourcentage de classes prédites correctement.

Les modèles que nous retiendrons finalement sont ceux qui présentent les meilleures performances en classification avec le minimum d'erreur.

4. Comportement des capteurs

Nous disposons des différents types de capteurs choisis pour la mesure des différents polluants, ces capteurs ont été testés et validés en laboratoire. Mais nous ne pouvons pas générer dans le laboratoire les mélanges gazeux rencontrés en situation réelle à cause de leur complexité. En conséquence que nous avons décidé d'étudier le comportement des réseaux de capteurs dans des situations réelles.

4.1. Etude des corrélations signaux des capteurs - polluants

Cette étude a été faite à partir des résultats de la campagne du site périurbain de Nogent, elle est réalisée à l'aide de PLS afin d'étudier et visualiser les corrélations entre les X variables contenant les données des capteurs et les Y variables contenant les informations sur la pollution. Nous avons utilisé les deux premières composantes principales (CP) résultantes, et avons tracé le graphe suivant en superposant les CP des X -variables (contenant les réponses des capteurs X_i et les variables ajoutées de l'humidité, de l'heure du jour et des jours ouverts) avec les CP des données physico-chimiques des 3 polluants NO, NO₂ et O₃ (voir figure 3-15). Dans cette étude nous utilisons les variations relatives de conductance ou de résistance (X_i) pour étudier les relations entre les réponses issues directement des capteurs. Dans ce graphe,

les variables sont représentées comme des points dans un espace à deux dimensions (2 CP). Dans un tel espace, la proximité entre les variables indique la relation entre eux : les points proches sont corrélés, les points diamétralement opposés sont anti-corrélés, et les points orthogonaux par rapport à l'origine sont indépendants. Notons que cette vision est très simplificatrice et qu'il est généralement nécessaire de la compléter par l'examen des composantes principales de rang supérieur à 2.

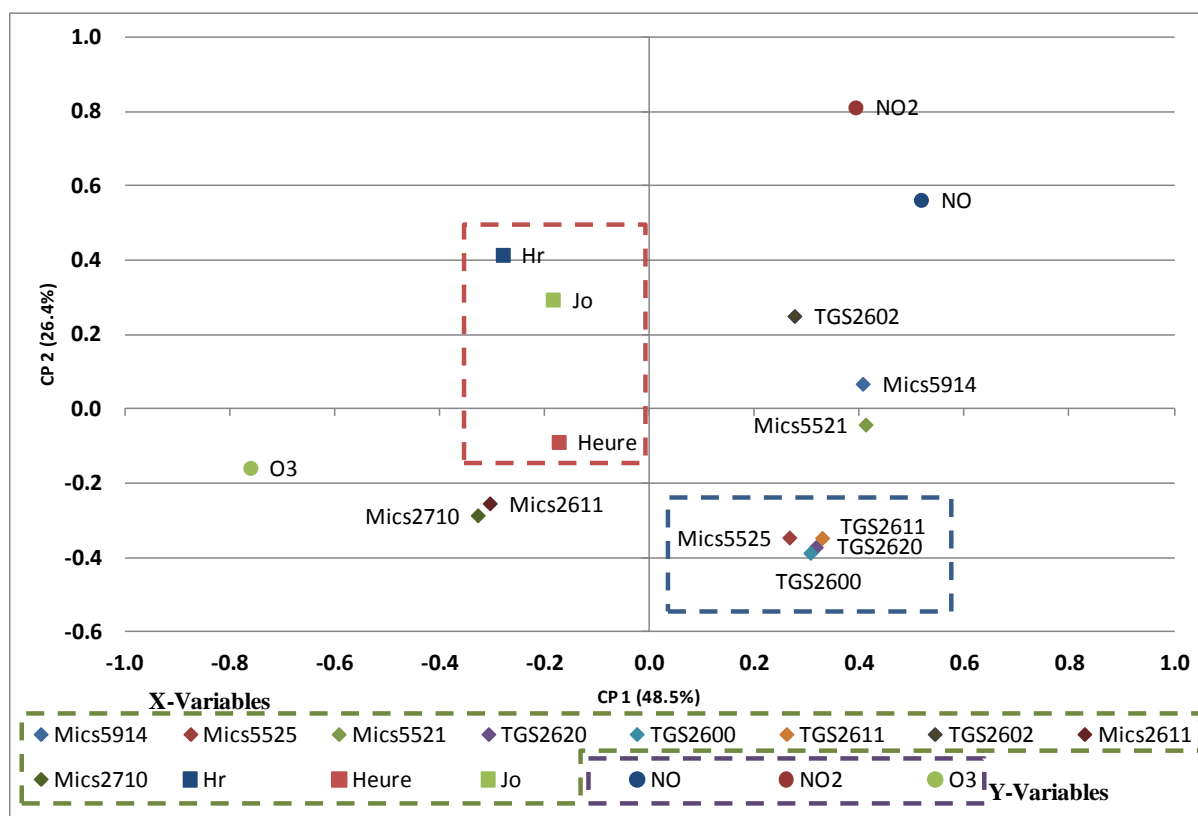


Figure 3-15 : Graphe résultant d'un calcul à base de PLS, réponses des capteurs issus de la première cellule à partir de la station de Nogent (Périurbaine).

Dans cet exemple, le point représentant l' O_3 est présent dans le quadrant opposé de celui représentant NO et NO_2 ce qui implique une anti-corrélation entre eux. Par ailleurs, la présence de NO et NO_2 dans le même quadrant avec une distance réduite entre eux implique la corrélation de ces deux variables. En effectuant le même raisonnement sur toutes les variables nous pouvons ainsi constater que :

- les oxydes d'azote (NO , NO_2) sont corrélés entre eux.
- l'O_3 est anti-corrélé avec les oxydes d'azotes, ces 2 premiers points confirment les observations faites au début de ce chapitre.
- Le capteur Mics2611 (capteur choisi pour la détection d' l'O_3) et le capteur Mics2710 (capteur choisi pour la détection de NO_2) ont des réponses corrélées avec l'O_3 et non pas avec NO_2 . Ce qui indique que le capteur Mics2710 est plus sélectif à l'O_3 que NO_2 , aspect important à prendre en considération pour les futures modélisations.
- Les capteurs Mics5521, TGS 2602 et Mics5914, capteurs choisis pour la détection des COV, ont des réponses corrélées avec la présence des oxydes d'azotes. Et comme nous l'avons déjà expliqué, NO provient des sources trafic, ces capteurs sont donc plutôt corrélés avec des différents polluants réducteurs qui proviennent de la même source de NO (comme le CO et les Hydrocarbures).
- Les réponses des capteurs TGS2600 (COV), TGS2620 (COV), TGS2611 (CH_4) et Mics5525 (CO) encadrées par le rectangle bleu sont indépendantes des polluants mesurés par analyses physico-chimiques. Comme le site est de type périurbain, les COV sont mal représentés, et donc ces capteurs mesurent des événements provenant de sources indépendantes. Ils sont donc riches en informations complémentaires.
- Les variables ajoutées (Hr, Heure, Jo) sont présentées avec des poids plus faibles que les autres variables, ce qui implique qu'elles contribuent d'une façon plus faible au système. Par contre, nous pouvons les corrélérer avec la présence de l'O_3 , c'est à dire la pollution photochimique présente durant la journée, comme montré dans la partie 3. Cette pollution photochimique est liée aux jours de la semaine mais aussi à la présence de pluie et à des différents événements météorologiques dont l'humidité est l'indicateur majeur.

- 3 capteurs, les TGS 2600, 2611 et 2620 semblent très proches les uns des autres.

L'examen de la 3eme composante principale nous montre que 2600 et 2620 donnent effectivement les mêmes informations, et qu'on devrait pouvoir faire l'économie d'un des deux capteurs.

Avec ces différentes observations, nous arrivons à mieux comprendre le comportement des capteurs dans une situation réelle, et pourrons ainsi mieux optimiser les méthodes de prédiction. Nous montrons d'autres études effectués sur les comportements des capteurs sur les sites de mesure dans l'annexe D page 209.

4.2. Comportement de type oxydant ou réducteur des capteurs

Lors des mesures réalisées sur les différents sites d'études, nous avons pu analyser les variations des conductances relatives de certains capteurs avec les évolutions des teneurs en polluants. Ainsi comme le montre la figure 3-16, les variations des conductances relatives des capteurs de COV, comme MICS 5521 et TGS 2602, sont relativement bien corrélées avec les variations de la concentration en NO₂ (seul polluant pour lequel l'étude des corrélations avec les concentrations est possible, les COV n'étant pas mesurés sur le site d'étude). Cependant, la présence de concentrations croissantes en NO₂ (gaz oxydant) dans l'air devrait se traduire par une variation négative de la conductance de ces capteurs or ces derniers présentent une variation de conductance positive qui ne peut être expliquée que par la présence de concentrations croissantes en gaz réducteurs vraisemblablement associés au NO₂. Or les gaz réducteurs réglementés (le monoxyde de carbone, le benzène,...) sont présents à de faibles concentrations individuelles et incapables de causer une telle réponse comme montré dans la figure 3-16 compte tenu de la faible sensibilité des capteurs à ces composés, ils ne peuvent donc pas expliquer individuellement le signal observé. Ainsi, nous constatons que les signaux mesurés sont probablement dus à la présence de gaz réducteurs non réglementés provenant de

la même source trafic que NO_2 , comme par exemple des alcanes provenant des émissions des échappements automobiles.

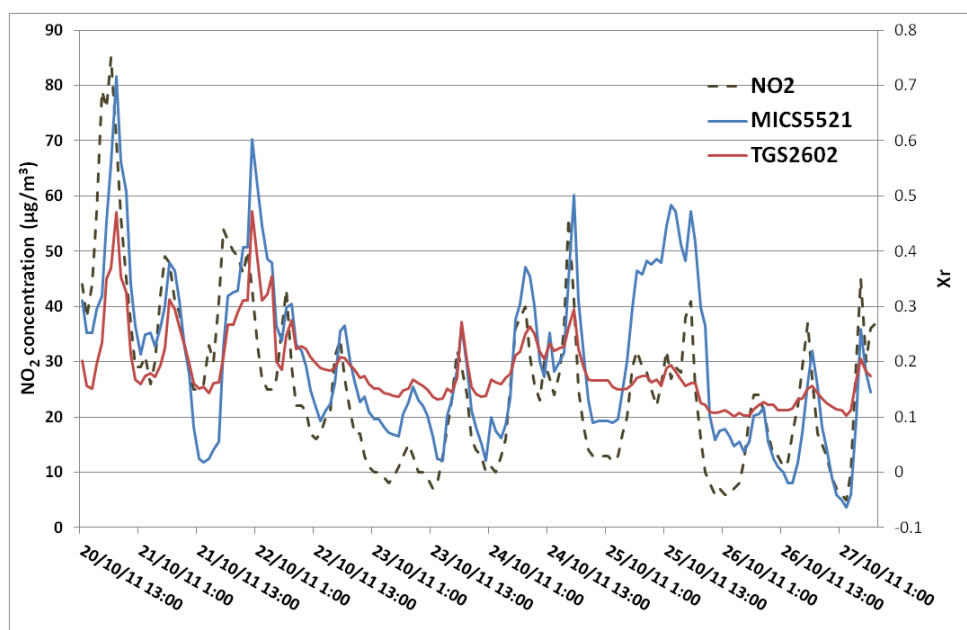


Figure 3-16 : Corrélation entre les capteurs de COV et la concentration de NO_2 dans le site de Valenciennes (Trafic).

4.3. Tentative de prédiction quantitative

Comme nous l'avons vu dans le chapitre I, les capteurs ou réseaux de capteurs sont généralement utilisés pour évaluer les concentrations de certaines espèces gazeuses (Tsujita, et al., 2005) (Helli, et al., 2004) (Negri, et al., 2001) (Sarry, et al., 1999) (Chaiboun, et al., 2007) (Szcsurek, et al., 2011) (Kamionka, et al., 2006)... Nous avons donc tenté ici de prédire directement les concentrations de certains polluants (NO , NO_2 et O_3) à partir des signaux des capteurs et autres variables explicitées plus haut, en utilisant les réseaux de neurones. Ce modèle obtenu a été optimisé avec une architecture ayant 12 entrées correspondant aux X variables et 3 sorties correspondantes aux 3 concentrations des polluants. Ce modèle possède une couche cachée avec 8 neurones. La figure 3-17 montre le résultat d'une autoprédiction d'un modèle calculé sur les données de la station périurbaine de Nogent à partir des réponses des capteurs de la cellule 1. Nous définissons par « autoprédiction », les résultats de l'application d'un modèle calculé sur les mêmes jeux de données utilisés pour construire ce

modèle (jeu d'apprentissage et de validation). Les bonnes performances des modèles d'autoprédiction est une condition nécessaire pour la pertinence de ces modèles, mais non suffisante pour la description de leurs performances globales et leur utilisation réelle, et donc leur capacité d'extrapolation.

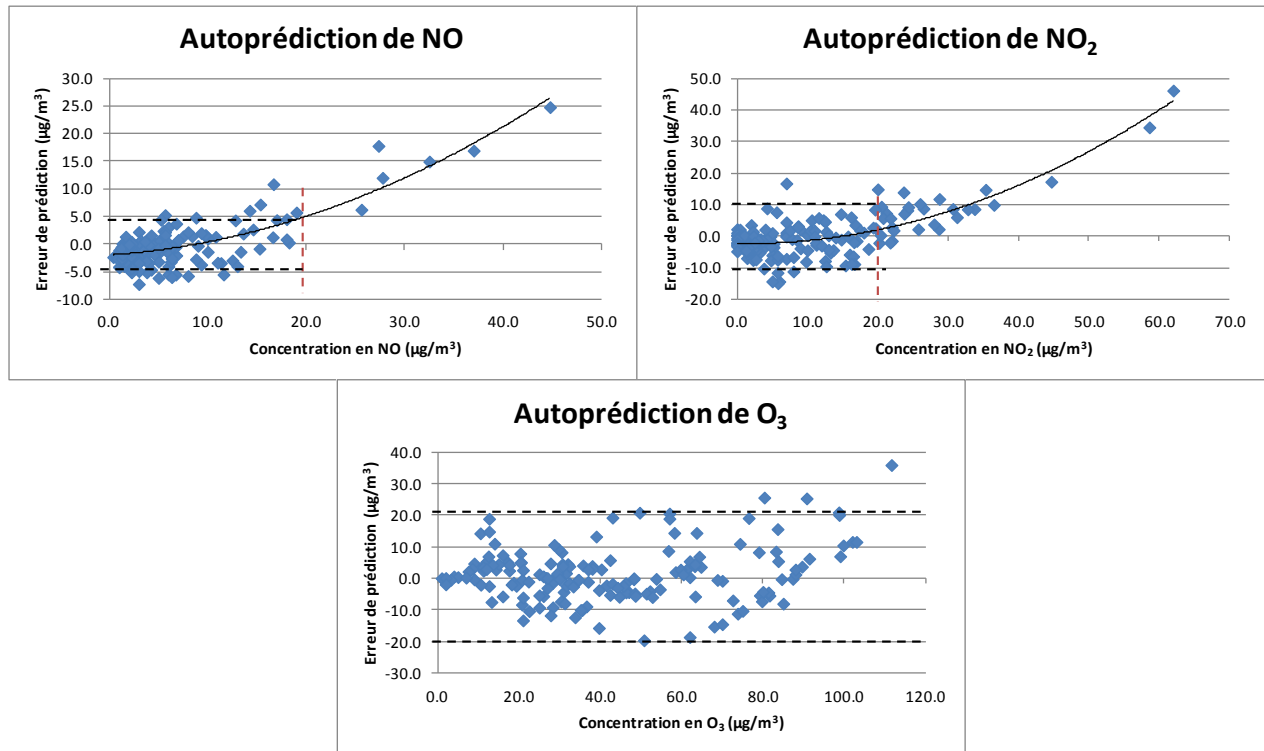


Figure 3-17 : Etudes quantitatives, erreurs de prédiction des concentrations des polluants à partir des réponses des capteurs et des variables ajoutées.

Dans cette figure nous traçons les erreurs de prédiction en fonction des concentrations réelles des polluants. Les performances de prédiction sont plutôt mauvaises. Pour des faibles concentrations qui ne dépassent pas $20\mu\text{g}/\text{m}^3$, l'erreur de prédictions est globalement limité dans un intervalle de $\pm 5 \mu\text{g}/\text{m}^3$ pour NO et $\pm 10 \mu\text{g}/\text{m}^3$ pour NO₂. Par contre, cette prédiction risque de surestimer les concentrations de NO et NO₂ pour des concentrations plus fortes que $20\mu\text{g}/\text{m}^3$, d'où les performances de prédiction sont plutôt mauvaises. Par contre pour l'O₃, nous pouvons constater que cette erreur est indépendante de la concentration. Dans ce dernier cas, nous notons que l'intervalle d'erreur est borné entre $[-20 \mu\text{g}/\text{m}^3 ; 20 \mu\text{g}/\text{m}^3]$ quelle que

soit la concentration. La prédiction est donc meilleure pour l'O₃ que celle de NO et NO₂ due à la haute sensibilité des capteurs à l'O₃.

Devant ces erreurs non négligeables obtenues et la difficulté à obtenir une haute précision avec une étude quantitative performante, nous avons orienté nos travaux plutôt vers une prédiction catégorielle. Ces difficultés de prédiction directe des concentrations sont par ailleurs liées au fait, observé plus haut, que les signaux des capteurs sont dus en bonne partie à des gaz non réglementés.

4.4. Prédiction catégorielle

Comme montré dans la partie 2 de ce chapitre, nous pouvons déterminer des types de pollution dans une zone urbaine à partir de la présence de trois polluants dans l'atmosphère (O₃, NO et NO₂). Donc, il est possible de catégoriser la pollution en 3 types : Urbain, Trafic et Photochimique. Des seuils doivent être choisis à partir des concentrations de ces polluants pour séparer ces catégories. Une étude de la distribution des concentrations (horaires) a été faite pour définir un seuil limite, au-delà duquel nous définissons la présence d'une catégorie de pollution.

Ces seuils ont été déterminés à partir de l'étude statistique d'une année d'historique de données sur les 7 sites d'étude (2010). Ils ne seront pas forcément pertinents pour d'autres sites (par exemple plus pollués), ou d'autres années. Ce choix a ici un caractère assez arbitraire et devra être adapté en fonction des sites et des objectifs recherchés.

4.4.1. Cas de la pollution trafic

Comme précisé antérieurement, un site trafic est globalement caractérisé par une forte concentration en NO et en NO₂ (voir partie 2 du chapitre III). La figure 3-18 montre la distribution annuelle en concentration de NO issue des sites de trafic étudiés, Roubaix et Valenciennes. L'étude de la distribution annuelle de NO sur les sites trafic montre un aspect

assez semblable. Ces deux distributions ont des médianes proches : $10 \mu\text{g}/\text{m}^3$ pour Valenciennes et $16 \mu\text{g}/\text{m}^3$ pour Roubaix. Les moyennes sont de $20 \mu\text{g}/\text{m}^3$ et $32 \mu\text{g}/\text{m}^3$ respectivement (voir tableau 3-2), avec un maximum d'occurrences situé entre 5 et $15 \mu\text{g}/\text{m}^3$ pour les deux distributions. Nous pouvons ainsi caractériser la présence d'une pollution trafic par l'observation de concentrations en NO supérieures à la valeur la plus haute des moyennes observées sur des sites trafic ainsi que l'absence de la pollution trafic par l'observation d'une concentration en NO inférieure à la valeur la plus basse des médianes observées sur des sites trafic.

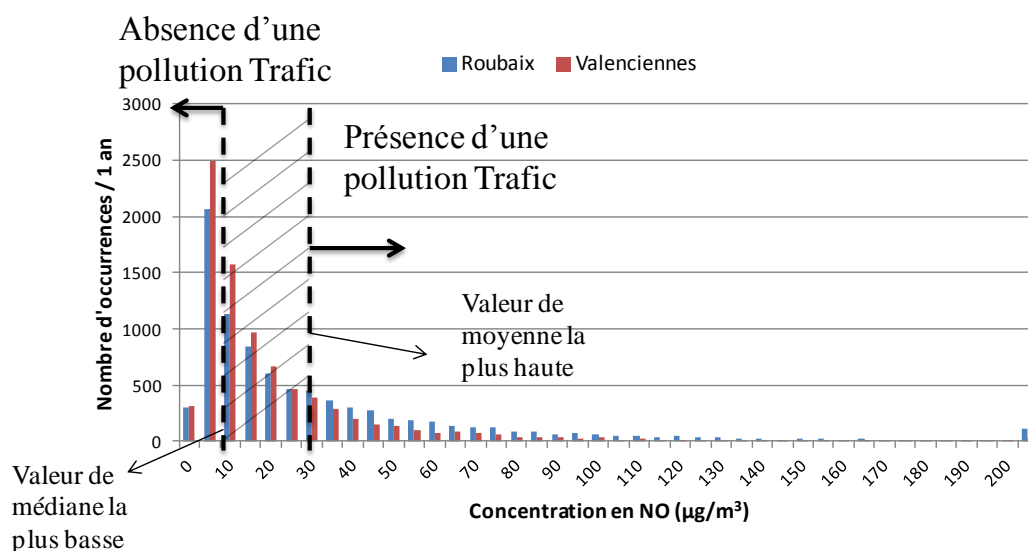


Figure 3-18 : Distribution annuelle en moyennes horaires des concentrations en NO dans les sites trafics (2010)

Nous avons choisi, pour des raisons d'approximations, un seuil de $30 \mu\text{g}/\text{m}^3$ pour définir la présence d'une source trafic. Cependant, il est important de noter que la sensibilité des modèles établis ultérieurement aux différents seuils pour les différents polluants sera également étudiée (cf § 6.1 « Seuils de catégorisation »).

4.4.2. Cas de la pollution urbaine

Nous avons effectué la même étude de distribution des concentrations en polluants sur des sites urbains, périurbains et trafic, avec, sur la figure 3-19 la distribution annuelle en NO_2

dans les sites urbains de Béthune, Douai et Creil, le site périurbain de Nogent et le site trafic de Roubaix. Comme pour le cas de NO, la distribution de NO₂ montre un aspect similaire dans les différentes stations urbaines et périurbaines, avec toujours une médiane proche de 20 µg/m³ et une moyenne proche de 25 µg/m³. Contrairement aux sites trafic, la moyenne de NO₂ peut monter jusqu'à 46 µg/m³ (voir tableau 3-2). Ainsi nous avons caractérisé, comme pour le cas précédent d'un site trafic, la présence d'une pollution urbaine par une présence en NO₂ de plus de 40 µg/m³. De la même façon que précédemment pour les sites « trafic », la sensibilité des modèles établis aux différents seuils pour les différents polluants sera également étudiée (cf 5.1.3. « Seuils de catégorisation »).

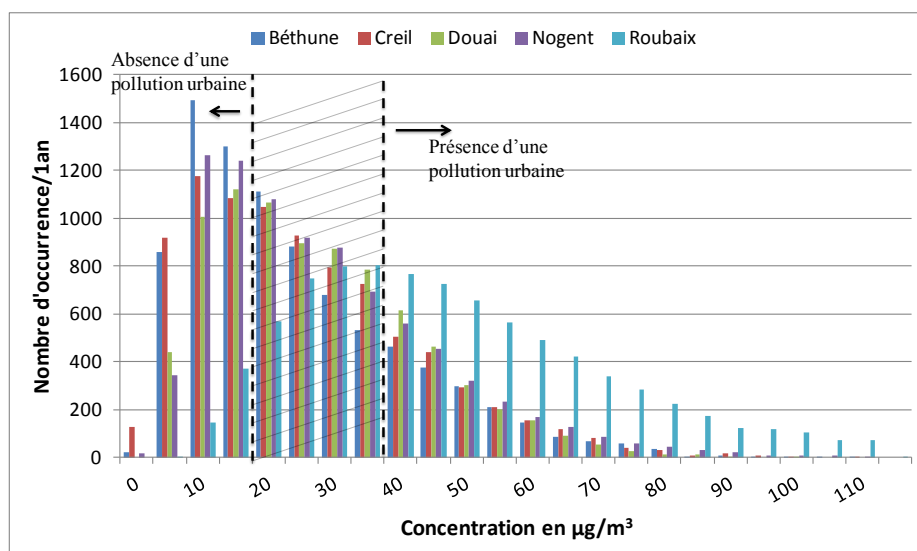


Figure 3-19 : Distribution annuelle en moyennes horaires des concentrations en NO₂ (2010)

4.4.3. Cas de la pollution photochimique :

Comme nous l'avons déjà montré, une pollution photochimique est caractérisée par la présence de concentrations relativement fortes en O₃ dans l'air. Comme pour les cas précédents, l'étude de la distribution annuelle de l'O₃ dans des sites urbains et périurbains montre des aspects semblables avec des valeurs moyennes de l'ordre de 40 µg/m³ et des valeurs médianes du même ordre de grandeur. Nous montrons sur la figure 3-20 la distribution d'O₃ durant une année de mesure (2010) sur les sites de Béthune, Creil, Douai et

Nogent. Nous pouvons alors définir un seuil supérieur de pollution photochimique au-delà de $50 \mu\text{g}/\text{m}^3$ d' O_3 pour des raisons d'approximations.

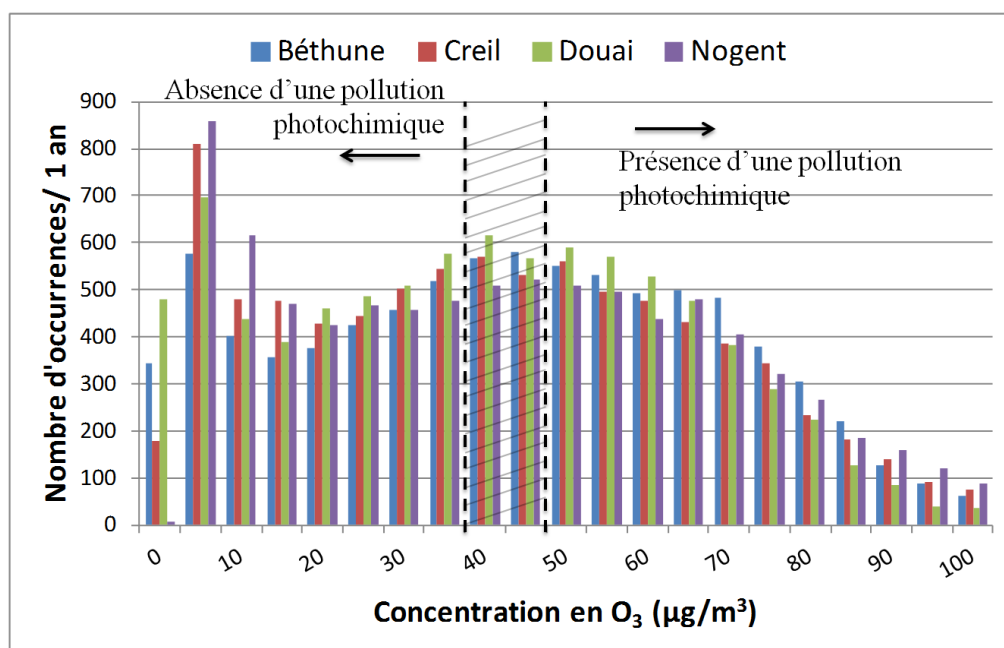


Figure 3-20 : Distribution annuelle en moyennes horaires des concentrations en O_3 (2010).

Ces seuils permettent de définir des catégories de pollutions, elles sont regroupées dans le tableau 3-3. Nous devons aussi introduire une catégorie dite « de pollution faible » qui définit les cas où les concentrations des trois polluants sont plus faibles que l'ensemble des seuils.

Tableau 3-3 : Tableau récapitulatif des seuils de catégorisation avec les types de pollution accordés.

	O_3	NO	NO_2
Seuils	$l = 50 \mu\text{g}/\text{m}^3$	$l' = 30 \mu\text{g}/\text{m}^3$	$l'' = 40 \mu\text{g}/\text{m}^3$
Type de pollution	Photochimique	Trafic	Urbain

4.4.4. Démarche de catégorisation :

En utilisant les seuils définis dans le tableau 3-3, nous pouvons classer la pollution de l'air en quatre catégories :

- Pollution trafic (T)
- Pollution urbain (U)
- Pollution photochimique (P)
- Pollution faible (F)

Comme l' O_3 et le NO sont fortement anticorrélés, il est possible d'illustrer cette catégorisation dans un plan comme montre la figure 3-21.

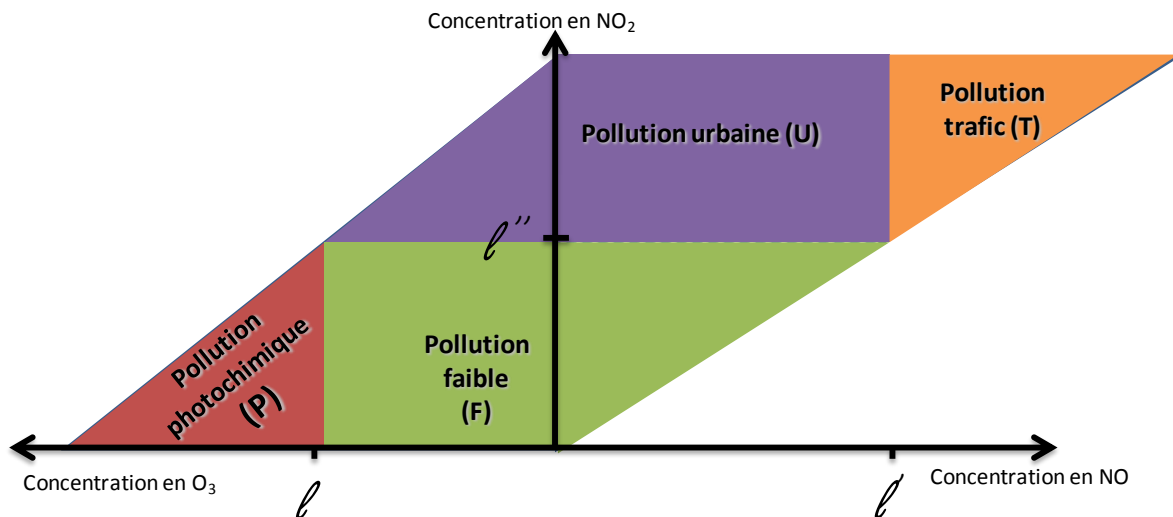


Figure 3-21 : Catégories de pollution en fonction des seuils de NO, NO₂, et O₃.

L'espace des classes représenté comme un losange montre la relation entre l' O_3 , NO, et NO₂ : une forte concentration de NO est accompagnée par une forte concentration de NO₂, et inversement, par rapport à l' O_3 . La pollution trafic peut être définie par la présence des concentrations élevées en NO et NO₂, la pollution photochimique peut être définie par la présence d'une concentration élevée en O_3 et supérieure au seuil l . La pollution urbaine sera le cas d'une forte concentration en NO₂. Nous pouvons toujours avoir des cas où la concentration en NO est beaucoup plus élevée que celle en NO₂ (c.f. figure 3-2, et annexe B),

spécialement dans des sites trafic, mais ceci n'affecte pas notre catégorisation, puisque la catégorie « trafic » est définie lorsque la concentration en NO dépasse le seuil de $30\mu\text{g}/\text{m}^3$.

Ces catégories peuvent être définies par des combinaisons booléennes entre les différentes concentrations des gaz comme montré dans le tableau 3-4.

Tableau 3-4 : détermination des catégories de pollution à partir des concentrations des polluants avec une classification booléenne.

Pollution catégorie	Catégorisation
Trafic (T)	$[NO]ET\overline{[O_3]} \approx [NO]$
Urbain (U)	$[NO_2]ET[NO]ET\overline{[O_3]}$
Photochimique (P)	$[O_3]$
Faible (F)	$\overline{[NO]}ET\overline{[NO_2]}ET\overline{[O_3]}$

Ces 4 catégories ne constituent pas toutes les combinaisons logiques mathématiquement possibles, car certaines ne peuvent pas exister, ainsi, par exemple une forte concentration en NO n'est jamais observée avec une faible concentration en NO₂, le cas $\overline{[NO_2]}ET[NO]$ n'existe donc pas.

Pour une telle catégorisation binaire, la valeur d'appartenance à une catégorie est 1 et la valeur de non appartenance à une catégorie est 0. Donc, les valeurs de sorties des modèles effectuant ce type de catégorisation sont des variables booléennes, étant 0 si la classe n'est pas détectée et 1 si elle est détectée. Nous avons calculé ainsi des modèles à bases de réseaux de neurones en cherchant à obtenir des classes définissant l'appartenance à une des catégories de pollution déjà présentées dans la figure 3-21 à partir des X-variables. Nous montrons dans le tableau 3-5 les taux de réussite de classification d'un modèle calculé sur les X-variables issues de la 3^{ème} cellule à partir des données de la station urbaine de Creil. Ce modèle est construit sur une architecture d'un réseau de 12-6-4. Rappelons que seules les performances de test sont représentatives des performances qu'on peut espérer avoir. Les performances d'apprentissage et de validation indiquent la modélisabilité du système.

Tableau 3-5: Taux de réussite de classification d'un modèle calculé selon une catégorisation binaire sur la station de Creil (Urbaine).

Classification	Trafic	Urbain	Photochimique	Faible	Total
Nb d'observations	0	8	102	56	166
Apprentissage	Pas de trafic	0/5 = 0%	55/55 = 100%	21/34 = 62%	76/94 = 84%
Validation	Pas de trafic	0/0 = 0%	22/24 = 92%	10/12 = 83%	32/36 = 83%
Test	Pas de trafic	0/3 = 0%	21/23 = 91%	1/10 = 10%	22/36 = 61%
Total	Pas de trafic	0/8 = 0%	98/102 = 96%	32/56 = 57%	130/166 = 78%

Nous rappelons que le taux de réussite de classification présenté dans le tableau 3-5 est le pourcentage des points dont la catégorie prédite (à partir des *X*-variables) est correcte selon la catégorie pointée par les données physico-chimiques correspondantes (les *Y*-variables). Le total indique la somme des performances obtenues par rapport au groupe désigné sur toutes les données, et le nombre d'observations indique le nombre de fois où les données physico-chimiques pointent à une classe de pollution donnée. Les performances pour les données d'apprentissage et de validation sont données pour mémoire, les performances réelles étant indiquées par les données dites de test.

Ainsi, lors de cette campagne, et à cause de l'épisode de pollution photochimique, la concentration de NO provenant de sources trafics était faible dans ce site et sur la période d'étude, donc la catégorisation des types de pollution à partir des données physico-chimiques des analyseurs automatiques n'a pas pu établir la présence d'une catégorie « trafic », et donc le nombre d'observations ayant la catégorie trafic était nul.

Lors de la modélisation à partir des données du système multi-capteurs, la pollution trafic n'a jamais été classifiée ce qui est en accord avec les données physico-chimiques. Ce n'est pas le cas de la catégorie « urbain », cette catégorie a été identifiée à partir des mesures des concentrations de polluants mesurées sur ce site, or chaque fois que les données physicochimiques pointent sur la catégorie « urbain », le modèle se trompe, ainsi aucune catégorie « urbain » n'a été détectée à partir des données issues des capteurs. Seule la

pollution photochimique prédomine, puisque les mesures ont été effectuées en période estivale. Ainsi nous trouvons qu'aucune catégorie urbaine n'a été correctement classifiée, contrairement à la catégorie photochimique qui a été justement classifiée à 96% sur toutes les données. Il apparaît également que le taux de réussite est de presque 84% sur toutes les données d'apprentissage et de validation, tandis qu'il n'est que de 61% sur les données de test qui sont des données inconnues par rapport au modèle établi. Nous remarquons alors la bonne détection d'une pollution « Photochimique », par contre, nous observons des mauvaises performances pour des pollutions « Faible » et « Urbaine ».

Nous pouvons alors constater que ce modèle traduit mal le comportement réel de la pollution sur le terrain. Il est possible d'affiner fortement cette classification en réglant de manière moins abrupte le basculement entre les classes aux frontières des seuils comme montré dans la figure 3-22 suivante. En effet, l'utilisation des fonctions booléennes pour les combinaisons des états des polluants implique que l'état =1 si la concentration du polluant est plus élevée que le niveau choisi et 0 si la concentration est plus faible. Une telle catégorisation peut créer des problèmes de frontières et de zones mal définies. Ainsi, un changement de $1\mu\text{g}/\text{m}^3$ peut brutalement faire basculer d'une catégorie à l'autre. Comme solution à cet effet de seuil, nous proposons d'utiliser la logique floue.

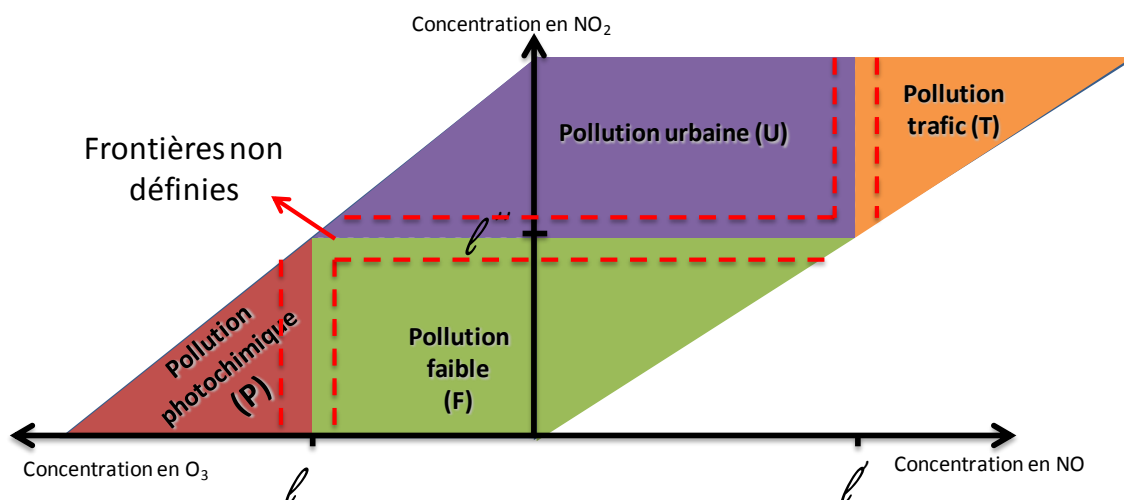
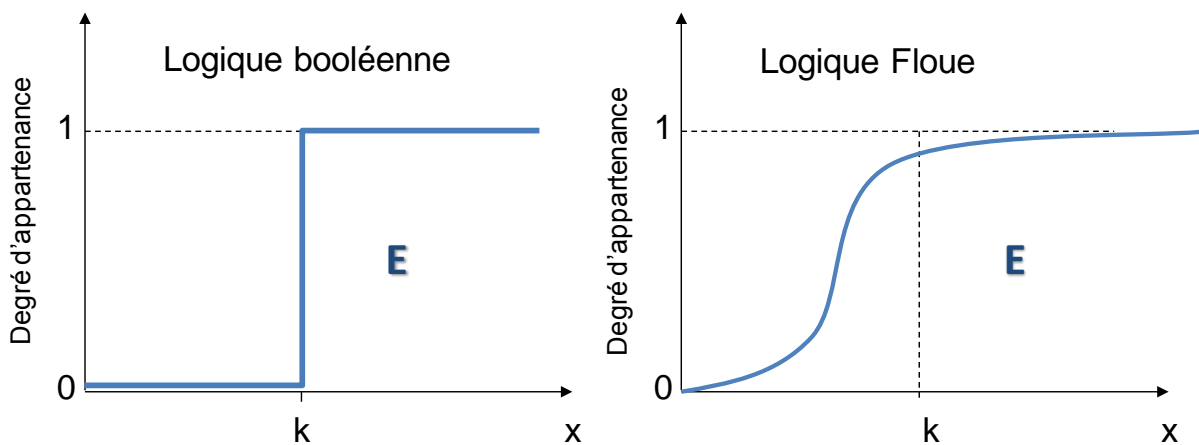


Figure 3-22 : Problèmes des frontières dans une classification booléenne.

4.5. Catégorisation avec la logique floue

La logique floue (« *fuzzy logic* ») est largement utilisée dans le contrôle intelligent, la détection des formes, et la classification des données. Elle est utilisée comme un substitut à la logique booléenne qui, de par sa limitation à deux états (vrai ou faux), présente une zone de passage non définie entre ces deux états comme montré dans la figure 3-22, et un effet de seuil très préjudiciable. La logique floue permet de modéliser les imperfections des données et imite dans une certaine mesure la flexibilité du raisonnement humain. Introduit en 1965 par Lotfi Zadeh (Zadeh, 1968) en se basant sur sa théorie mathématique des ensembles flous, la logique floue introduit la notion de degré d'appartenance permettant à une condition d'être dans un état autre que vrai ou faux contrairement à la logique booléenne. Elle offre une flexibilité et permet de prendre en compte les inexactitudes et les incertitudes d'un système et résout le problème de l'effet de seuil. La figure 3-23 montre l'exemple d'une équation logique: $x > k \rightarrow x \in E$ en utilisant la logique booléenne et la logique floue.



Le changement brusque du degré d'appartenance observé avec la logique booléenne est substitué par une augmentation continue de la valeur du degré d'appartenance avec la logique floue. Par conséquent, l'application de la logique floue aux concentrations de polluants peut permettre une catégorisation de la pollution ne causant pas de problème d'effet

de seuil. Ainsi nous passons d'une approche de catégorisation avec des variables booléennes (0 ou 1) vers une approche de catégorisation avec des variables continues qui varient entre 0 et 1. Ces nouvelles variables seront les degrés d'appartenance aux catégories de pollutions choisies. Ainsi la catégorie ayant le degré d'appartenance le plus élevé sera considérée comme la catégorie prédominante (le type de pollution prédominant).

4.5.1. Choix d'une fonction non linéaire

Afin d'appliquer la logique floue à notre catégorisation des types de pollution, des fonctions d'appartenance non linéaires sont nécessaires pour décrire l'évolution des concentrations de NO, NO₂ et O₃ dans l'atmosphère. Ces fonctions doivent être des fonctions non linéaires, monotonement croissantes et bornées entre 0 et 1.

Nous avons ainsi choisi comme fonction de base une fonction couramment utilisée en logique floue et définie par :

$$f(x) = \frac{1}{2} \left(1 + \tanh(2p(x - a)) \right) \quad (2)$$

Cette fonction a les propriétés suivantes (voir figure 3-24) :

- Non linéaire
- Bornée entre 0 et 1
- Monotone croissante.
- $f(a) = 0.5$ (point d'inflexion)
- p : défini la pente de la fonction pour $f(x) = 0.5$.

Cette fonction peut aussi être définie par deux seuils comme montré dans la figure 3-24, ce qui correspond mieux à nos besoins car on pourra faire le lien avec les seuils de concentrations établis pour séparer les classes de pollution.

- $F(I_{inf}) = 0.12 = q$
- $F(I_{sup}) = 0.88 = 1-q$

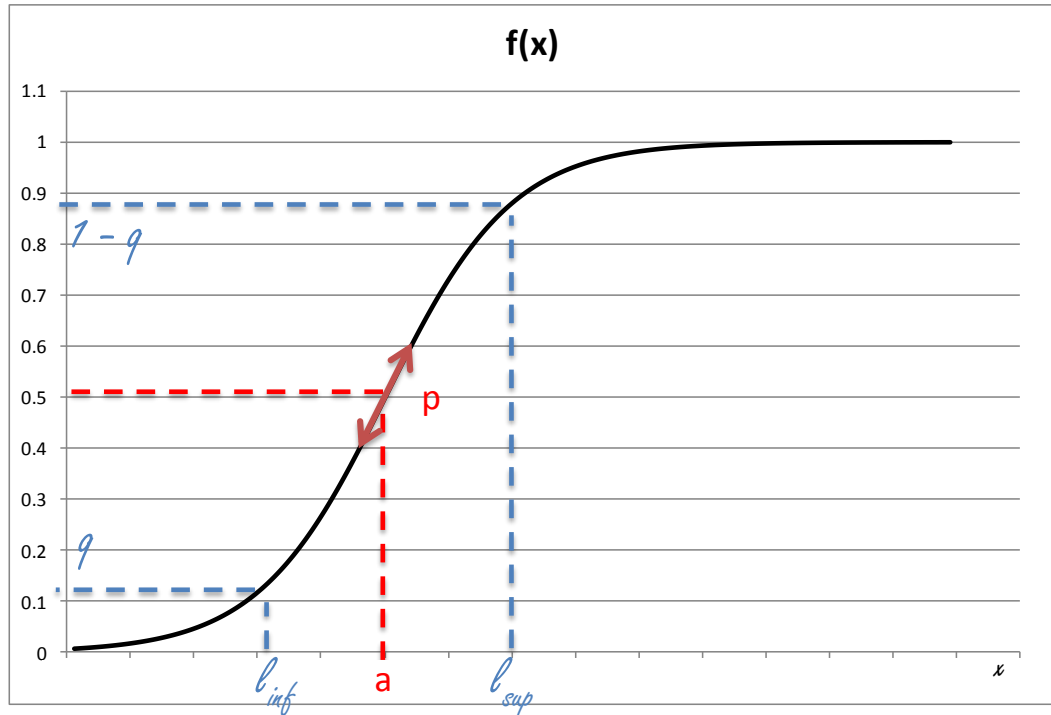


Figure 3-24 : La fonction non linéaire choisie.

En introduisant ces deux seuils dans l'équation (2), nous pouvons écrire :

$$F(x) = \frac{1}{2} \left(1 + \tanh \left(\ln \left(\frac{1-q}{q} \right) \left(\frac{x - L_{0.5}}{Dl} \right) \right) \right) \quad (3)$$

Avec :

- $L_{0.5} = \left(\frac{l_{inf} + l_{sup}}{2} \right) = a$
- $Dl = \frac{l_{sup} - l_{inf}}{2} = 2p$

Concrètement, l_{sup} est la valeur de concentration au-delà de laquelle la catégorie de pollution est très probable, et l_{inf} la valeur en deçà de laquelle la catégorie de pollution est très improbable. Ainsi si nous nous basons sur les seuils qui ont été présentés dans le tableau 3-3 et dans les figures 3-18, 3-19 et 3-20, nous supposons que quand les concentrations des polluants sont plus faibles que les valeurs de médianes annuelles (calculées sur 2010), ces concentrations ne peuvent pas contribuer à une catégorie de pollution véritable, ce sont les l_{inf} , et inversement quand les concentrations de polluants dépassent leurs moyennes annuelles (calculées sur 2010) nous considérons que ces concentrations peuvent contribuer à une

pollution d'où les seuils présentés dans le tableau 3-3 deviennent les seuils supérieurs (l_{sup}). Ainsi nous pouvons ré-établir et les nouveaux seuils de pollutions en se basant sur les valeurs de la partie 4.4, comme montré dans le tableau 3-6.

Tableau 3-6 : Tableau récapitulatif des seuils de catégorisation

Seuils	NO	NO ₂	O ₃
l_{inf}	15 µg/m ³	20 µg/m ³	30 µg/m ³
l_{sup}	30 µg/m ³	40 µg/m ³	50 µg/m ³

4.5.2. Application de la logique floue à la catégorisation des pollutions

Les relations de catégorisation ainsi obtenues pour calculer le degré d'appartenance de chaque catégorie à partir des fonctions d'appartenance des polluants utilisent les seuils du tableau 3-6 et se basent sur les relations montrées dans le tableau 3-4. Ces fonctions floues sont le résultat des combinaisons avec des opérateurs flous (Zadeh, 1968) entre les fonctions d'appartenance de chaque polluant comme montré dans le tableau 3-7.

Tableau 3-7 : Les opérateurs logiques écrits en logique booléennes et logique floue (Zadeh, 1968).

Opérateur	Equation booléenne	Equation floue
Et	$X \text{ et } Y$	$\min\{f(X), g(Y)\} = f(X) * g(Y)$
Ou	$X \text{ ou } Y$	$\max\{f(X), g(Y)\} = f(X) + g(Y)$
Non	\bar{X}	$1 - f(X)$

La fonction utilisée peut être ainsi écrite comme ceci :

$$F_i(x) = \frac{1}{2} \left(1 + \tanh \left(\ln \left(\frac{0.88}{0.12} \right) \left(\frac{2x - l_{inf_i} - l_{sup_i}}{l_{sup_i} - l_{inf_i}} \right) \right) \right) \quad (4)$$

$$\begin{aligned} i &= 1 \text{ pour NO} \\ i &= 2 \text{ pour NO}_2 \\ i &= 3 \text{ pour O}_3 \end{aligned}$$

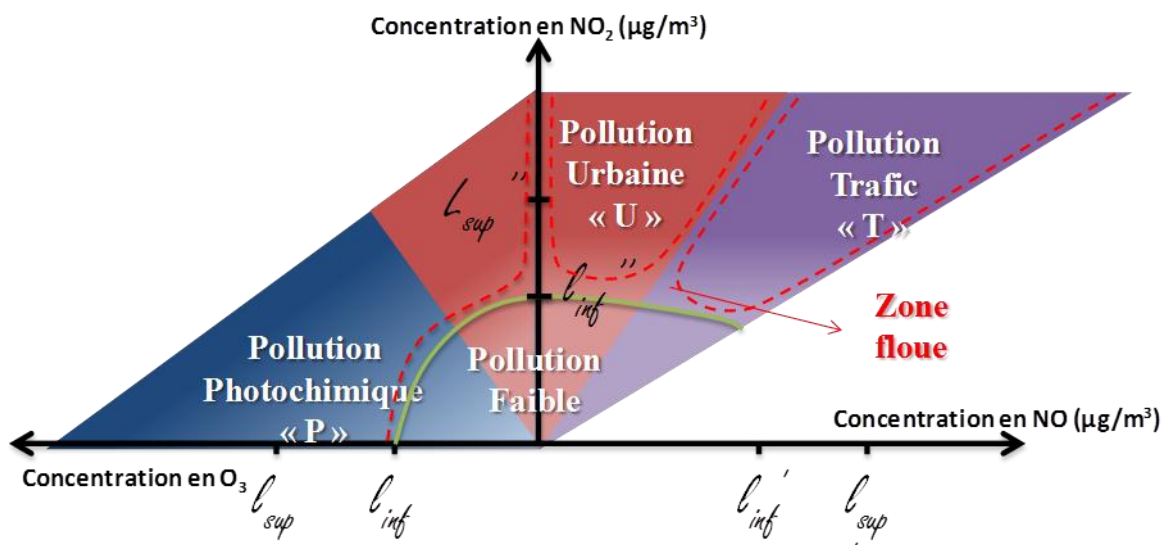
Le tableau 3-7 des équations de calcul des catégories de pollution peut ainsi être réécrit dans le tableau 3-8 :

Tableau 3-8 : détermination des degrés d'appartenance des catégories des pollutions à partir des concentrations des polluants.

Catégorie pollution	Equation booléenne	Equation floue
Trafic (T)	$[NO] * \overline{[O_3]}$	$F_1(NO) * (1 - F_3(O_3))$
Urbain (U)	$[NO_2] * \overline{[NO]} * \overline{[O_3]}$	$F_2(NO_2) * (1 - F_1(NO)) * (1 - F_3(O_3))$
Photochimique (P)	$[O_3]$	$F_3(O_3)$
Faible (F)	$\overline{[NO]} * \overline{[NO_2]} * \overline{[O_3]}$	$(1 - F_1(NO_2)) * (1 - F_2(NO)) * (1 - F_3(O_3))$

Pour utiliser cette nouvelle méthode de catégorisation, il faut calculer les degrés d'appartenance de chaque catégorie de pollution à partir des valeurs horaires de concentrations de chaque polluant. Ces valeurs seront utilisées comme des variables cibles (Y-variables) pour le calcul des modèles à l'aide du réseau de neurones. Une classe de pollution sera considérée comme prépondérante, lorsqu'elle présente le degré d'appartenance le plus élevé. Le taux de réussite, est calculé comme le taux des prédictions correctes des classes de pollution.

La figure 3-25 illustre la nouvelle catégorisation des polluants en utilisant la logique floue, les zones non définies sont des zones dont on connaît désormais la contribution sous formes de degré d'appartenance à chaque catégorie, ainsi, il est toujours possible de calculer la pollution prépondérante dans cette zone.

**Figure 3-25 : Illustration de la catégorisation des pollutions en se basant sur la concentration des polluants**

4.5.3. Résultat de catégorisation

Reprenons l'exemple de la station de Creil présenté dans la partie 5.4.4 en appliquant cette nouvelle méthode de catégorisation. Nous montrons ici la démarche de catégorisation des différentes catégories de pollutions sur ce site urbain. L'évolution de la concentration de ces polluants est montrée figure 3-26-A. Nous pouvons observer l'évolution de la concentration en O_3 durant les heures des journées ainsi que des faibles pointes associées au trafic automobile du matin montrées par les pics en NO et NO_2 . Les fonctions d'appartenance de chaque polluant sont alors calculées à partir des valeurs de concentrations et de l'équation (4) en utilisant les seuils définis par le tableau 3-6, Ces fonctions sont montrées figure 3-26-B.

Nous remarquons que la forme d'évolution de ces fonctions est proche de celle des concentrations des polluants (figure 3-26-A), par contre, les amplitudes des variations sont réduites et bornées entre 0 et 1, ce qui est normal puisque nous cherchons à faire une catégorisation et pas une quantification. Nous pouvons ainsi calculer les degrés d'appartenance prédits (les Y -variables) à partir des données des concentrations des 3 polluants, qui sont montrés figure 3-26-C. Nous retrouvons ainsi les épisodes de pollutions photochimiques correspondant aux concentrations élevées en O_3 tracés en verts, les faibles pics de pollution trafic le matin correspondant aux concentrations en NO et NO_2 émises par des sources trafic. Et finalement, une pollution urbaine est observable durant les nuits dues aux concentrations élevées en NO_2 .

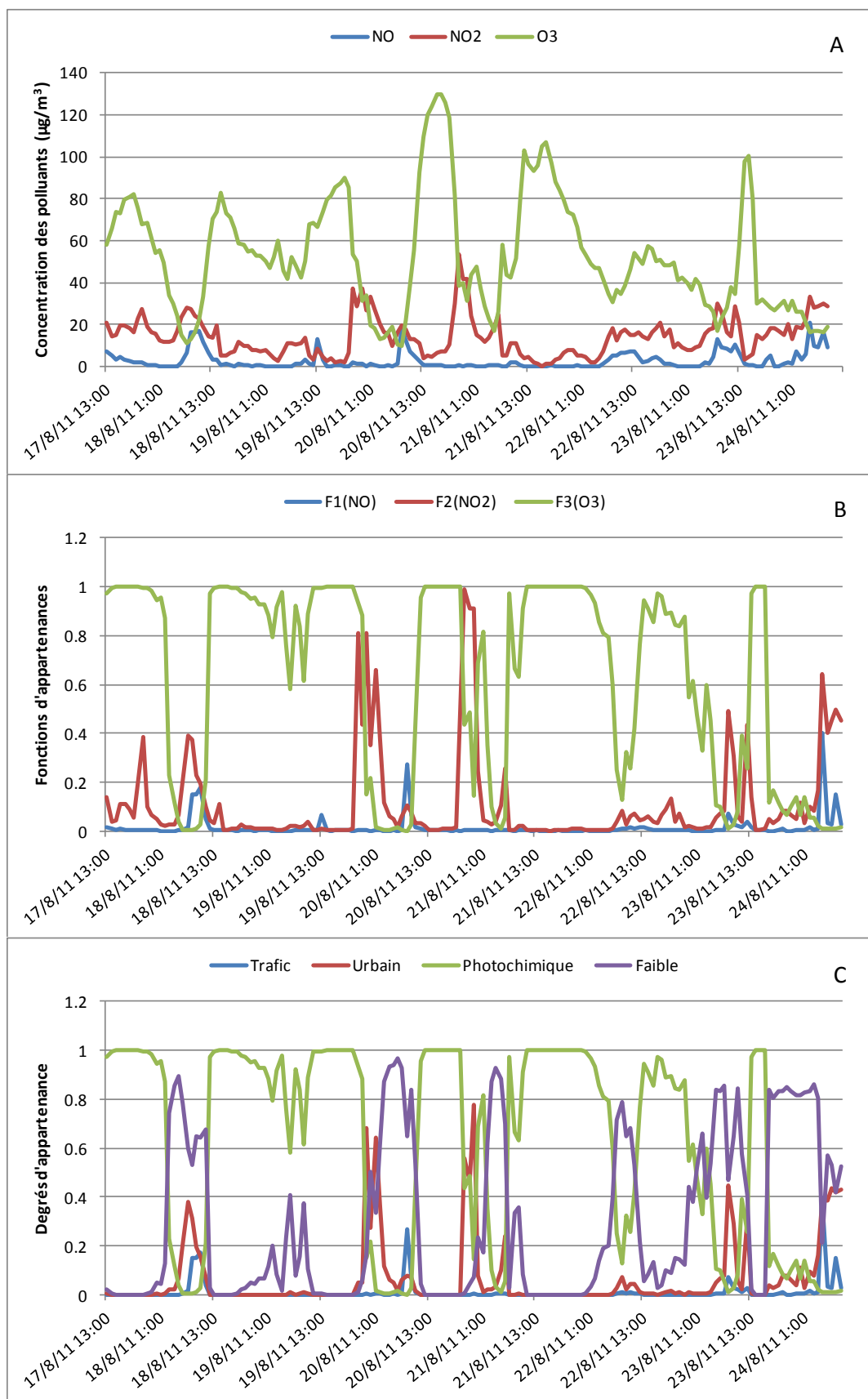


Figure 3-26 : démarches de catégorisations des types de pollution à partir de la campagne de mesure du site de Creil.

Après modélisation par réseau de neurones nous effectuons une comparaison entre catégorisation à partir des données physicochimique (*Y*-variables) et la catégorisation à partir des *X*-variables (*Y*-variables prédites) pour calculer un taux de réussite de classification. Il s'agit ici « d'autoprédiction », les problèmes d'extrapolation seront vus plus loin. Le modèle utilisé a une architecture en 12-4-4. Les résultats de prédiction de ce modèle sont montrés dans le tableau 3-9. Nous effectuons une comparaison avec les résultats obtenus par la méthode de catégorisation binaire étudiée dans la partie 4.4.4.

Tableau 3-9 : Tableau représentant les taux de réussite des modèles calculé pour prédire des catégories booléennes et floues.

Catégories	Classification	Trafic	Urbain	Photochimique	Faible	Total
	Nb d'observations	0	13	92	61	166
Booléen Y-variables booléennes	Apprentissage	Pas de trafic	0/5 = 0%	55/55 = 100%	21/34 = 62%	76/94 = 84%
	Validation	Pas de trafic	0/0 = 100%	22/24 = 92%	10/12 = 83%	32/36 = 83%
	Test	Pas de trafic	0/3 = 0%	21/23 = 91%	1/10 = 10%	22/36 = 61%
	Total (%)	Pas de trafic	0/8 = 0%	98/102 = 96%	32/56 = 57%	130/166 = 78%
Floue Y-variables floue	Apprentissage	Pas de trafic	5/5 = 100%	54/55 = 98%	30/34 = 88%	89/94 = 95%
	Validation	Pas de trafic	0/0 = 100%	23/24 = 96%	7/12 = 58%	30/36 = 83%
	Test	Pas de trafic	3/3 = 100%	23/23 = 100%	9/10 = 90%	35/36 = 97%
	Total (%)	Pas de trafic	8/8 = 100%	100/102 = 98%	46/58 = 82%	154/166 = 93%

Ainsi il est possible de voir que sur le site de Creil sur la période où les mesures ont été faites, aucune pollution de type trafic n'a été identifiée par les deux méthodes de catégorisation (booléenne et floue). Par contre le type urbain n'a jamais pu être identifié lors de l'utilisation de la logique booléenne alors que la réussite de classification est de 100% avec utilisation de la logique floue. De la même façon pour les types de pollution photochimique et faible, l'utilisation de la logique floue amène une nette amélioration des taux de réussite de classification, en utilisant des réseaux de neurones moins complexes qu'auparavant. La pollution prédominante est celle qui possède le degré d'appartenance le plus élevé d'une catégorie.

Cette méthode de prédiction est déjà très intéressante mais souffre de son côté 100% catégoriel. Elle n'apporte aucune information quantitative concernant le degré de pollution, même si la catégorie « pollution faible » introduite, constitue en quelque sorte une quantification binaire de l'intensité de pollution.

4.6. Prédiction mixte catégorielle et indicielle

Afin d'introduire une notion quantitative à la catégorisation des types de pollution, nous avons décidé de nous inspirer, sans le reproduire à l'identique, de l'indice de pollution « ATMO » utilisé dans la communication grand public par les AASQA (voir chp1 partie 2.4.) en France. L'indice ATMO qui est le maximum de 4 indices mineurs calculés à partir des concentrations de 4 polluants (SO_2 , O_3 , NO_2 et PM_{10}), est utilisé pour informer le public sur la qualité de l'air ambiant dans une région. Chaque indice varie de 1 à 10, 1 étant le cas d'une très faible pollution et 10 le seuil maximal d'alerte définie pour un polluant à partir des directives européennes (chapitre I partie 2). Dans notre cas, l'indice de pollution choisi « IP » sera basé sur 3 indices mineurs calculés à partir des concentrations disponibles de polluants, soit NO , NO_2 et O_3 . Notre indice IP sera aussi défini comme étant le maximum des 3 indices mineurs. Les indices d' O_3 et de NO_2 peuvent être calculés de la même façon que celle définie par ATMO. Par contre, comme NO n'apparaît pas dans l'indice ATMO, un nouvel indice a été créé pour décrire sa présence.

Nous savons que NO est caractéristique de la pollution trafic, et que pour une telle pollution les valeurs de NO et NO_2 sont corrélées. Cet indice est alors estimé de façon à maintenir la même distribution que celle de l'indice de NO_2 par une simple règle de 3 entre la distribution annuelle des concentrations de NO_2 et celle de NO . Cette distribution sera calculée sur l'historique des données de l'année 2010. La figure 3-27 suivante montre le nombre d'occurrence des concentrations de NO et NO_2 dans toutes les stations étudiées. Donc nous pouvons en déduire le tableau 3-10 suivant.

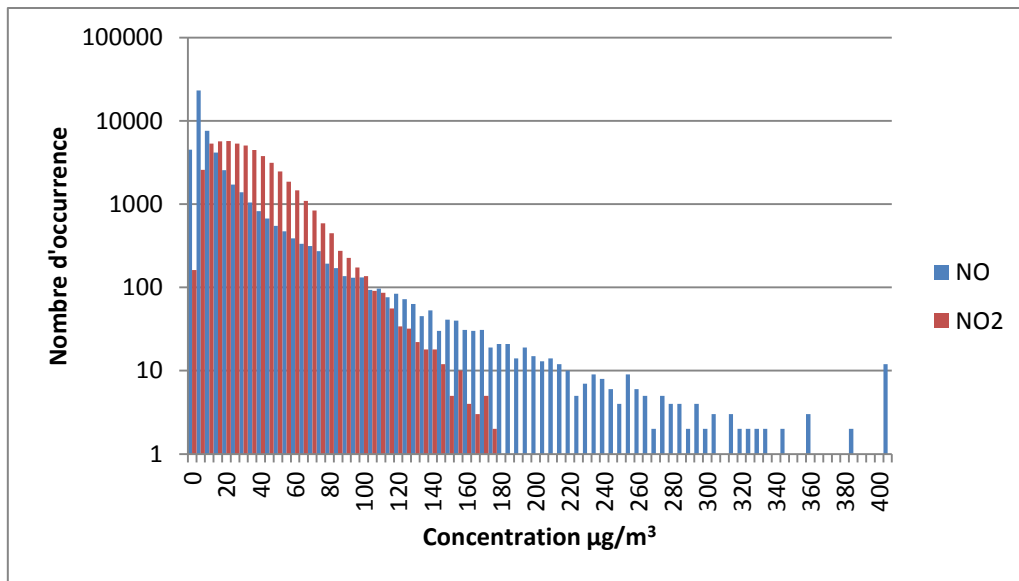


Figure 3-27 : fréquence d'occurrence en moyenne horaire des concentrations de NO et NO₂ durant l'année 2010.

Tableau 3-10 : Indice de pollution développé pour aider à obtenir une catégorisation mixte.

Indice	NO ($\mu\text{g}/\text{m}^3$)	NO ₂ ($\mu\text{g}/\text{m}^3$)	O ₃ ($\mu\text{g}/\text{m}^3$)
1	0-9	0-29	0-29
2	10-34	30-54	30-54
3	35-74	55-84	55-79
4	75-154	85-109	80-104
5	160-239	110-134	105-129
6	239-314	135-164	130-149
7	315-379	165-199	150-179
8	380-439	200-274	180-209
9	440-499	275-399	210-239
10	≥ 500	≥ 400	≥ 240

Cet indice IP, qui nous permettra d'obtenir des informations sur le degré de pollution, sera calculé, de la même façon que les degrés d'appartenance, à partir des concentrations issues des analyseurs pour l'apprentissage, puis à l'aide du réseau de neurones lors de la prédiction. Les catégories ne seront alors plus que 3 puisque la catégorie « pollution faible » n'a plus de raison d'être.

Mais cet indice est un indice discrétisé, ce qui limite les informations et ajoute une erreur de discrétisation sur nos modélisations. Nous avons donc décidé d'établir des équations continues pour affiner cet indice. Durant nos campagnes de mesures, nous n'avons jamais rencontré des cas où l'indice ATMO dépasse la valeur 6, ce qui explique pourquoi dans notre calcul l'indice maximal est 7. La figure 3-28 montre les équations obtenues jusqu'à l'indice 7. Nous remarquons que les indices de NO₂ et O₃ obtenus à partir des indices ATMO sont quasi-linéaires, ce qui n'est pas le cas pour l'indice de NO calculé.

De cette manière, la nouvelle catégorisation prendra la forme mixte présentée sur la figure 3-29. La catégorie de pollution faible, est substituée par la valeur de l'indice de pollution IP.

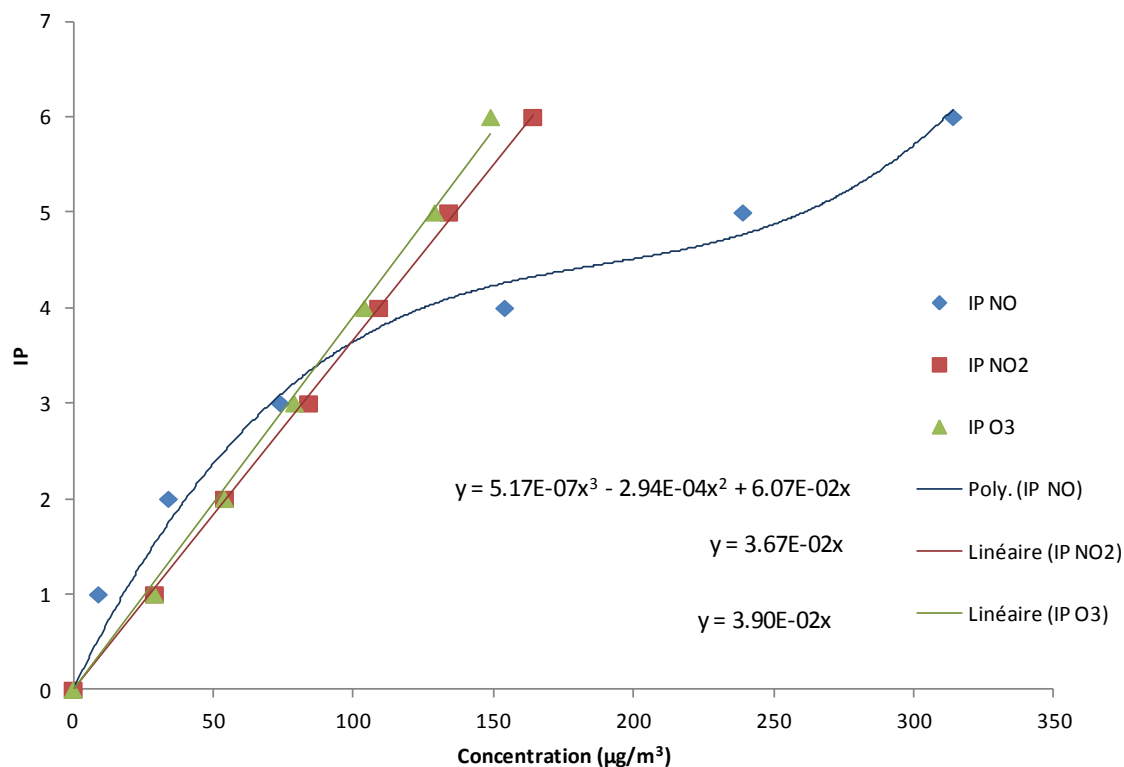


Figure 3-28 : Les équations modélisant les 3 indices mineurs.

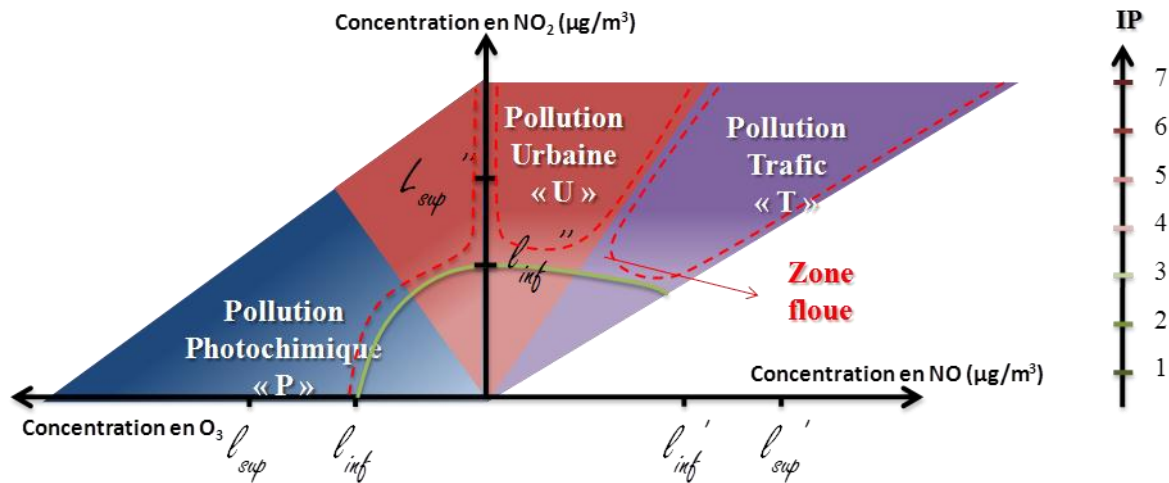


Figure 3-29 : Nouvelle catégorisation mixte, la catégorie de la pollution faible est substituée par l'indice de pollution.

La partie suivante de ce chapitre, sera consacrée à l'optimisation des modèles de prédiction et à leur portabilité et utilisera cette approche mixte, catégorielle et indicielle.

5. Résultats de prédiction mixte

Nous présentons dans cette partie les travaux effectués pour l'optimisation des performances des modèles, ainsi que ceux validant les choix déjà présentés dans les parties précédentes. Dans la suite nous ne présentons que les performances correspondantes à la partie de test des modèles calculés pour montrer leurs robustesses vis à vis des nouvelles valeurs inconnues.

5.1. Optimisation des modèles

Nous avons débuté nos modélisations en étudiant la modélisabilité du système. Les réponses des différentes cellules ont été utilisées pour prédire les catégories de pollutions dans chaque site ainsi que l'indice de pollution « IP ».

Dans un premier temps, et afin d'optimiser la réalisation des modèles, il a fallu s'assurer de la pertinence des différents choix concernant les variables, les traitements et les seuils évoqués plus haut.

5.1.1. Normalisation et standardisation des sensibilités

Deux critères de performances des modèles développés ont été utilisés :

- Le RMSEP (*Root Mean Squared Error of Prediction*) qui quantifie l'écart entre les cibles et les sorties du réseau.
$$RMSEP = \sqrt{\frac{\sum (Y - Y')^2}{n}}$$
- Taux de classification réussie, c'est le pourcentage de classes prédites correctement.

Le tableau 3-11 montre l'effet du choix des variables des réponses de capteurs sur les performances des modèles calculés. Les modèles suivants sont calculés pour prédire les différentes catégories de pollution ainsi que l'indice de pollution. Dans la recherche d'une modélisation où les critères ne sont pas applicables dans les périodes d'épisodes photochimique, nous avons choisi pour cette modélisation la station de Béthune (urbaine). Trois modèles sont calculés, un en utilisant les réponses des capteurs non standardisées, un autre en utilisant les réponses des capteurs standardisées, et un troisième en utilisant les réponses standardisées et normalisées.

Tableau 3-11 : performances des modèles calculés sur le site de Béthune.

Réseau de capteurs	Modèle par cellule					
	Performances avec les réponses des capteurs non standardisées (X_i)		Performances avec des réponses des capteurs standardisées (S_i)		Performances avec des réponses des capteurs standardisées et normalisées (Z_i)	
	Réussite	RMSEP-IP	Réussite	RMSEP-IP	Réussite	RMSEP-IP
Cellule 1	80 %	0.3	85%	0.07	92%	0.03
Cellule 2	75 %	0.29	82%	0.08	98%	0.02
Cellule 3	78 %	0.3	86%	0.08	98 %	0.03

L'utilisation des réponses des capteurs standardisées en sensibilité pour la modélisation apporte une meilleure performance aux modèles, ce qui confirme notre choix d'utiliser que les réponses normalisées à partir des réponses standardisées en sensibilité dans la modélisation.

5.1.2. Choix des variables

Le choix des variables contrôle le taux des informations fournies au modèle lors de sa construction. Enrichir un modèle avec des variables supplémentaires peut augmenter son taux de réussite, ou au contraire complexifier inutilement le modèle. Le tableau 3-12 montre un test de modélisation effectué sur la station de Nogent (périurbaine). Deux modèles ont été calculés, le premier en n'utilisant que les réponses normalisées des capteurs, le deuxième en incluant les variables supplémentaires choisies qui sont : l'humidité relative de l'air, l'heure du jour, et les jours ouvrés.

Tableau 3-12 : performances comparées des modèles calculés sur le site périurbain de Nogent.

Réseau de capteurs	Modèles par cellule			
	Performances avec les réponses des capteurs normalisées seulement		Performances avec les variables supplémentaires	
	Réussite	RMSEP-IP	Réussite	RMSEP-IP
Cellule 1	89 %	0.16	90.5%	0.03
Cellule 2	87%	0.2	91%	0.013
Cellule 3	85%	0.12	89%	0.014

L'ajout de variables supplémentaires améliore nettement les taux de réussite. Confirmant ainsi l'utilité d'enrichir les données d'entrées par ces 3 variables.

5.1.3. Seuils de catégorisation

Le choix des valeurs des seuils de catégorisation peut influencer les performances des modèles, et l'appartenance à une catégorie. Le tableau 3-13 suivant montre l'effet de la variation des seuils de catégorisation. Cet exemple de calcul est appliqué à la station de Nogent, en changeant les seuils hauts l_1 et bas l_2 .

Les modèles correspondant aux valeurs choisies initialement par défaut (tableau 3-6), montrés sur la première ligne, présentent les meilleures performances. En effet, le fait de varier les seuils en diminuant le seuil bas (l_{inf}) ou augmentant le seuil haut (l_{sup}) va élargir ou rétrécir la

zone floue d'une catégorie mais les critères de performance restent sensiblement voisines sauf pour la cellule 2 qui semble la plus sensible.

Tableau 3-13 : effet de la variation des seuils sur les critères de performance du modèle.

Station	NO ($\mu\text{g}/\text{m}^3$)		NO ₂ ($\mu\text{g}/\text{m}^3$)		O ₃ ($\mu\text{g}/\text{m}^3$)		Cell1		Cell2		Cell3	
	I_{inf}	I_{sup}	I_{inf}	I_{sup}	I_{inf}	I_{sup}	Taux réussite	RMSEP- IP	Taux réussite	RMSEP- IP	Taux réussite	RMSEP- IP
Nogent	15	30	20	40	30	50	90.5%	0.03	91%	0.013	89%	0.014
	10	30	15	40	20	50	92.5%	0.03	89%	0.015	86%	0.019
	20	40	30	50	60	50	87%	0.03	86%	0.021	85%	0.019

Le choix des seuils est important pour avoir des résultats riches en informations, mais il est toutefois guère possible d'optimiser les performances de prédiction en jouant sur les seuils. Car il est possible à la limite d'avoir des taux de succès de 100% (mais sur une seule catégorie) en choisissant des seuils inf. et sup. respectivement petit et grand.

5.2. Modèles mono-cellule et mono-site, modèles locales.

Après optimisation du choix des variables et des seuils, nous avons cherché à comparer les différences de performances entre les trois cellules, en réalisant un modèle différent pour chaque cellule ET chaque site, donc dans un premier temps sans chercher à extrapoler. Le tableau 3-14 montre les performances obtenues avec des modèles calculés donc par cellule et par station. Les valeurs montrées dans ce tableau sont obtenues à partir de données utilisées pour le test, ces performances sont ainsi obtenues sur des données considérées comme inconnues par rapport à leurs modèles.

Tableau 3-14 : Performances des modèles mono-cellules et mono-sites.

		Cellule 1		Cellule 2		Cellule 3	
Site	Type	Taux réussite	RMSEP IP	Taux réussite	RMSEP IP	Taux réussite	RMSEP IP
Creil	Urbaine	87%	0.04	90%	0.03	89%	0.05
Nogent	Périurbaine	90%	0.03	91%	0.01	89%	0.01
Saint-Amand	Périurbaine	89%	0.03	87%	0.03	88%	0.03
Valenciennes	Trafic	91%	0.02	90%	0.01	89%	0.02
Béthune	Urbaine	91%	0.03	98%	0.02	98%	0.03
Douai	Urbaine	91%	0.01	92%	0.01	93%	0.02

Durant ces modélisations, un réseau à une seule couche cachée contenant 4 neurones (12-4-4) est optimal. Toutes les performances données dans les tableaux 3-9 à 3-14 sont obtenues avec des réseaux à 4 neurones cachés (12-4-4). Les performances des 3 cellules sont proches les unes des autres. Nous détaillons ces résultats dans le cas de la cellule 3 pour le site de Béthune dans la figure 3-30.

Les réponses des capteurs enregistrées à partir de la cellule 3 durant cette campagne ont été utilisées pour prédire les degrés d'appartenances à des catégories de pollution. Le graphe A présente les degrés d'appartenance calculés à partir des concentrations de polluants tandis que le graphe B montre les valeurs prédites pour la troisième cellule. Ce modèle calculé à un taux de réussite de classification de 98% avec un RMSEP de l'indice de pollution de 0.03 comme déjà montré dans le tableau 3-14. La figure 3-31 montre les valeurs réelles et prédites d'IP sur le même site et la même cellule.

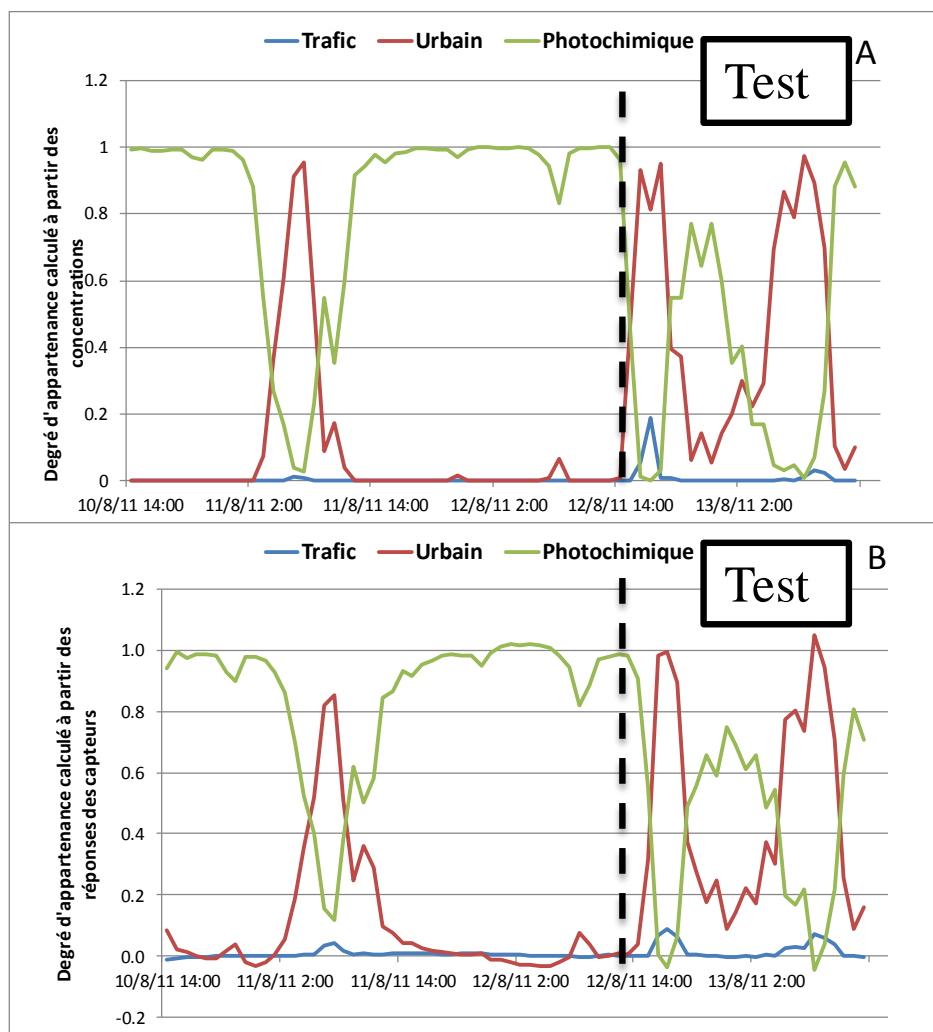


Figure 3-30 : Catégories prédites à partir des données de la 3eme cellule issue de la station de Béthune

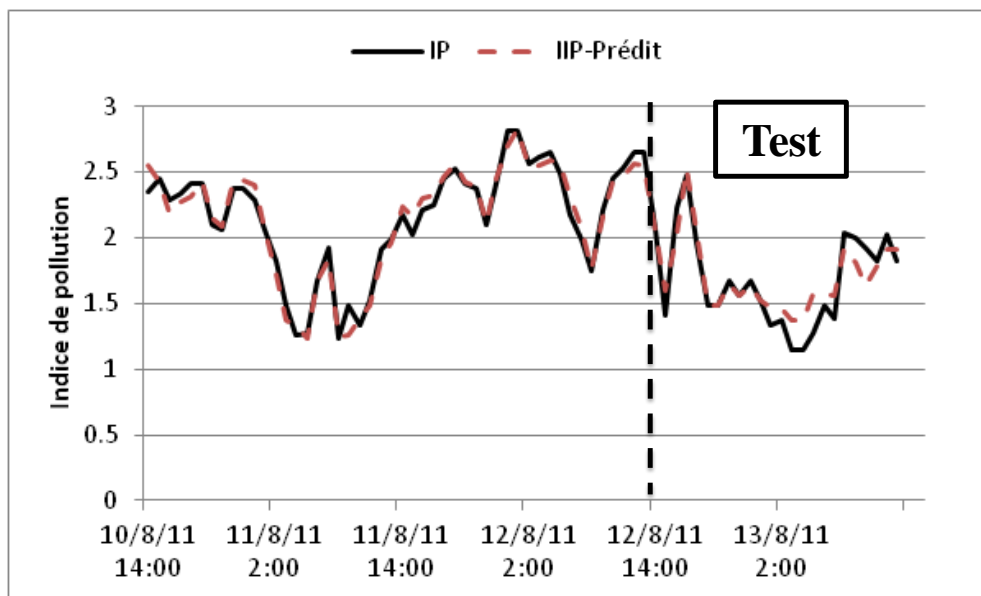


Figure 3-31 : Indice de pollution à partir des données de la 3eme cellule issue de la station de Béthune comparé à l'indice prédit

5.3. Portabilité des modèles

La portabilité d'un modèle consiste à savoir dans quel domaine de validité le modèle donné peut être utilisé. Ainsi, une modélisation étant une opération longue et délicate, il peut être intéressant d'avoir un modèle « universel » valable pour toutes les cellules. On peut de même envisager un modèle valable sur plusieurs sites de même catégorie ou de catégories différentes.

5.3.1. Modèles mono-cellule et inter-site

Nous avons testé la portabilité des modèles en utilisant un modèle construit à partir des données d'une cellule sur **une** station pour prédire les degrés d'appartenances à partir de la même cellule mais dans **une autre** station de même type. Nous avons alors trouvé qu'un modèle établi à partir des données recueillies sur **une** station n'est pas capable de modéliser les données d'autres stations, même si c'est une station de même type. Par exemple le modèle calculé sur Nogent (périurbain) à partir de la cellule 1 avec un taux de réussite de classification de 90 % avec $RMSEP-IP = 0.03$, en l'appliquant sur le site de Saint Amand (périurbain) nous obtiendrons une réussite de classification de 25% et un $RMSEP-IP = 0.1$. La figure 3-32 montre l'écart de performance du modèle calculé à partir des données de la station de Nogent et appliqué à la station de Saint Amand.

Ce mauvais comportement montre l'impossibilité d'extrapolation d'un modèle mono-site sur une station inconnue, chaque site ayant probablement des particularités. Un modèle intersites devra donc être calculé à partir d'une base de données intégrant plusieurs sites, de différents types.

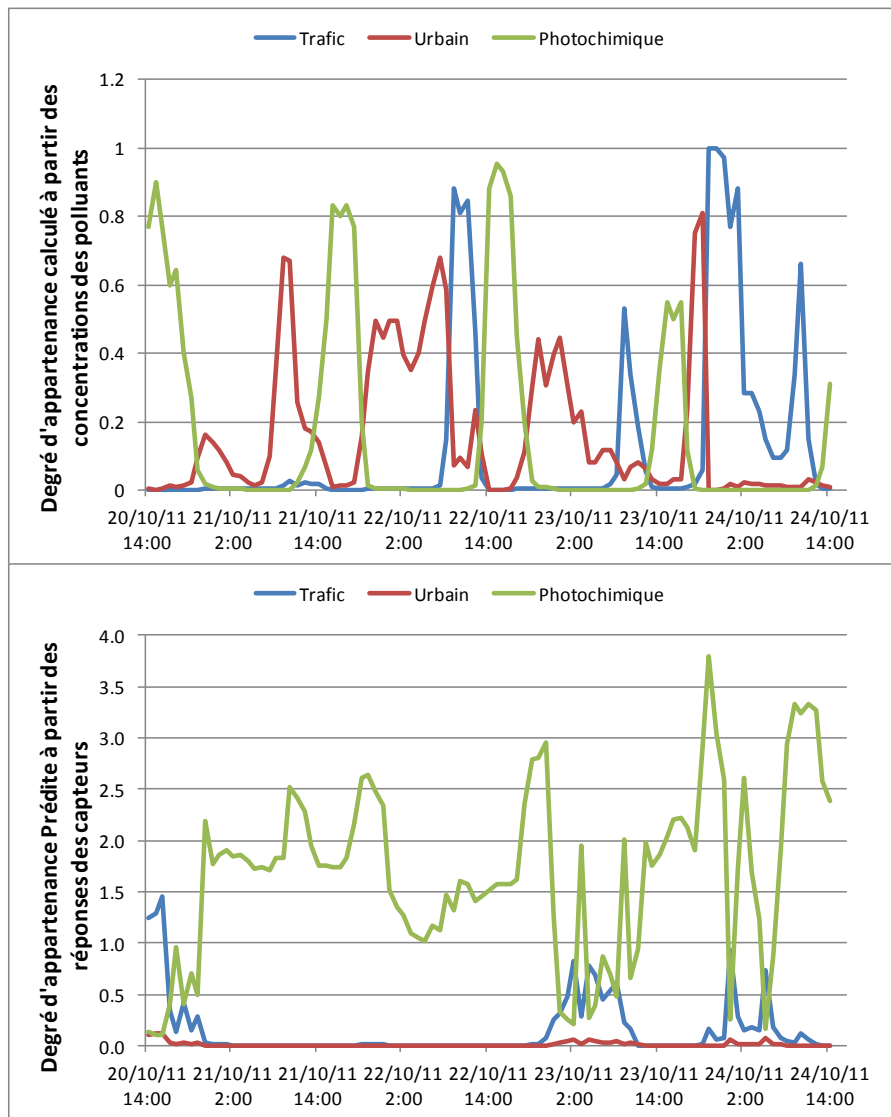


Figure 3-32 : modèle calculé sur Nogent et appliqué sur Saint Amant (cellule-1)

5.3.2. Modèles mono-cellule et multi-site

Nous avons ainsi essayé de développer un modèle capable de prendre en compte les mesures et événements des différentes stations. Ces modèles sont toujours construits cellule par cellule (modèles mono-cellules). Les groupes d'apprentissage et de validation ont été pris à partir de toutes les stations (de tous les types) excepte la station de Béthune dont les données ont été utilisées pour le test, donc ces données ne rentreront pas dans le calcul du modèle. Ainsi nous obtiendrons un modèle construit à partir de différentes stations et sera appliqué sur une station inconnue. Le tableau 3-15 présente les performances de ces modèles face à une station inconnue (station de Béthune).

Tableau 3-15 : Performance des modèles multi-station en face d'une station inconnue (Béthune)

	Cellule 1		Cellule 2		Cellule 3	
Données	Classification	RMSEP-IP	Classification	RMSEP-IP	Classification	RMSEP-IP
Béthune (modèle local - tableau 3-14)	91%	0.03	98%	0.02	98%	0.03
Béthune (modèle multi- station)	86	0.09	84	0.06	86	0.08

Le taux de réussite pour la station de Béthune est d'environ 85% contre 91% à 98% en « modèle local ». Nous remarquons aussi une légère dégradation du RMSEP de l'indice de pollution, cependant ce résultat reste satisfaisant.

Pour développer des modèles capables de prédire les catégories de pollutions sur des sites inconnus, il faut disposer d'une large base de données provenant des différents types de stations mesurées, y compris bien sûr du type de la station inconnue, ce qui permet au modèle de reconnaître les particularités de ce type des stations. Cela se traduit alors par une plus grande robustesse du modèle (domaine de validité plus grand), mais bien sûr des performances moins élevées que pour un modèle spécifique à un site.

5.3.3. Modèles inter-cellules.

Nous avons vu qu'un modèle calculé à partir de données issues d'un seul site est difficilement capable de réaliser la prédiction de bonne qualité sur d'autres. Nous avons vu qu'en élargissant la base de données et en y incorporant les données de plusieurs sites nous arrivons à prédire, avec une erreur acceptable, des cas inconnus. Nous poursuivons cette démarche « d'extrapolation » afin de modéliser différentes cellules dans le but d'obtenir un modèle valable pour toutes les cellules.

Nous avons alors étudié la faisabilité d'un modèle calculé sur deux cellules. Prenons l'exemple d'un modèle calculé sur la station de Creil (Urbain) en utilisant les données des cellules 1 et 3. Ce modèle calculé aboutit à un taux de réussite de classification de 85% sur

ces deux cellules avec un RMSEP de 0.024. Ainsi un modèle construit sur plusieurs cellules sera capable de reconnaître les différentes caractéristiques de ces cellules. Par contre de ce modèle appliqué à la cellule 2, ne présente un taux de réussite de classification que de 64% avec un RMSEP de 0.3. Cela signifie :

- qu'un modèle calculé sur plusieurs cellules arrive à reconnaître les particularités des différentes cellules utilisées lors de la construction, ce qui enrichit ce modèle.
 - que la performance obtenue en appliquant un modèle calculé sur deux cellules vers une cellule inconnue est nettement moins bon.
 - nous remarquons néanmoins pour ce dernier cas un meilleur comportement qu'avec un modèle mono-cellule calculé à partir une cellule et appliqué sur une autre « inconnue ».
- La figure 3-33 montre un bilan de telles modélisations. Un exemple d'une telle démarche est présenté dans le tableau 3-16.

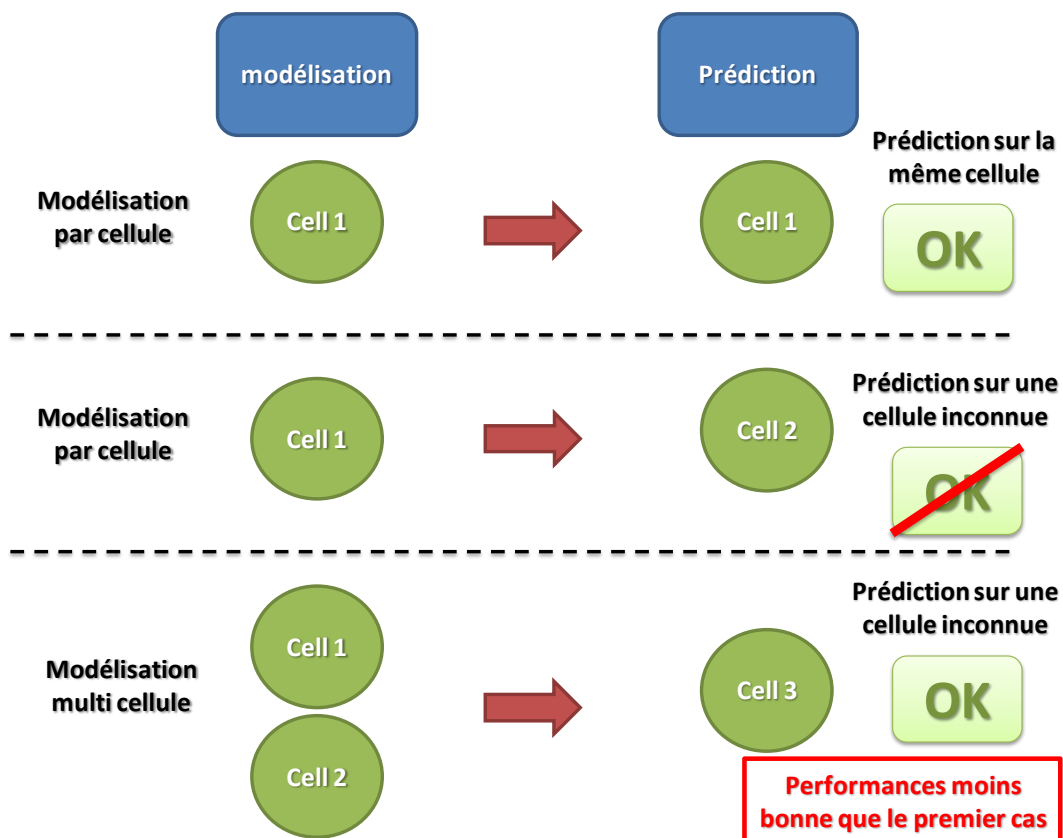


Figure 3-33 : Bilan des performances de différents modèles inter-cellules.

Nous considérons qu'une prédiction est satisfaisante (présenté comme OK sur la figure 3-33) quand les taux de réussite dépassent 80% pour le groupe de test, et que tous les catégories soit prédites avec un taux de réussite au moins égale à 70%. Nous trouvons aussi qu'un RMSEP-IP dépassant 0,3 est insatisfaisant.

Tableau 3-16 : Exemple des performances des modèles inter cellules calculées sur la station de Nogent (périurbaine).

Modèle	Prédiction	Taux de réussite	RMSEP-IP
Cellule 1	Cellule 1	90%	0.03
Cellule 2	Cellule 2	91%	0.01
Cellule 3	Cellule 3	89%	0.01
Cellule 1	Cellule 2	48%	0.2
Cellule 1+2	Cellule 1+2	83%	0.06
Cellule 1+ 2	Cellule3	72%	0.1

La non reproductibilité de la réponse des capteurs nuit aux performances des modèles sur des cellules inconnues. Même après avoir effectué une standardisation des sensibilités des capteurs, sensée corriger ce problème, le problème est atténué mais subsiste. Nous avons effectivement vu dans le chapitre II que la standardisation des sensibilités n'est pas parfaite puisqu'elle ne porte que sur un seul gaz cible.

Néanmoins, l'utilisation de plusieurs cellules lors de la modélisation permet au modèle de reconnaître différentes situations, et de tenir compte partiellement de cette non reproductibilité des modèles. Ainsi un modèle calculé sur plusieurs cellules est toujours mieux adapté à l'application à des cellules « inconnues » qu'un modèle calculé sur une seule cellule. Comme déjà vu pour le cas inter-sites, nous avons montré que plus la base de données d'un modèle est riche, et plus ce dernier sera capable de modéliser des cas inconnus et sera donc universel, mais au prix de performances moins bonnes que celles d'un modèle plus spécifique.

Nous pouvons donc conclure qu'un modèle calculé à base de réseaux de neurones ne peut pas modéliser des situations inconnues (extrapolation).

5.4. Amélioration du critère de performance

Le calcul du taux de réussite de classification, tel qu'il a été calculé ici, est effectué sur l'ensemble des données, y compris celles qui présentent un indice de pollution faible. Pour ces données, les faibles valeurs de concentration peuvent introduire une erreur de classification en raison de la confusion possible entre les différentes catégories de pollution, alors que cette dernière est quasiment inexistante. Par conséquent, si nous considérons comme succès toute classification dès lors que l'IP (calculé et réel) correspondant est au-dessous d'un seuil fixé, il est possible d'éviter cette confusion. Cela est équivalent à rétablir automatiquement la catégorie « *F = Faible pollution* » évoquée plus haut, dès que l'IP est faible. Le tableau 3-17 reprend l'exemple du modèle calculé sur toutes les stations et présenté dans la partie 5.3.2., nous montrons dans ce tableau les performances totales incluant la partie test sur Béthune de ce modèle en variant le seuil de IP.

Tableau 3-17 : le taux de prédiction du modèle selon le seuil d'IP choisi.

	IP ≥ 0	IP ≥ 1	IP ≥ 2	IP ≥ 3	IP ≥ 4	IP ≥ 5
Cellule 1	80%	83%	97%	99%	100%	100%
Cellule 2	82%	83%	95%	98%	100%	100%
Cellule 3	82%	82.1%	91%	98%	100%	100%

Si on incorpore les données ayant une basse IP dans le calcul du taux de classification, la catégorie de pollution identifiée risque d'être facilement confondue avec un autre type de pollution. A l'opposé, un seuil trop élevé élimine toute classification. Là encore, le choix de ce seuil est un compromis relativement arbitraire entre performance apparente et richesse de l'information obtenue. Par contre avec l'indice $IP \geq 2$ nos résultats sont nettement améliorés.

5.5. Sensibilité des modèles aux variables

La sensibilité des modèles aux variables est un indicateur montrant l'importance des X-variables, et donc notamment des capteurs lors de la modélisation. Cet indicateur de sensibilité est automatiquement calculé par Statistica lors de la construction d'un modèle.

Au cours des modélisations sur les stations urbaine et périurbaine, les capteurs MICS2611 et MICS2710 dédiés à la détection des gaz oxydants présentent les sensibilités les plus élevées (4.7 et 3.8 respectivement) et les variables de l'heure du jour et des jours ouvrés présentent les sensibilités les plus basses (1.02 et 0.98) comme le montre la figure 3-34.

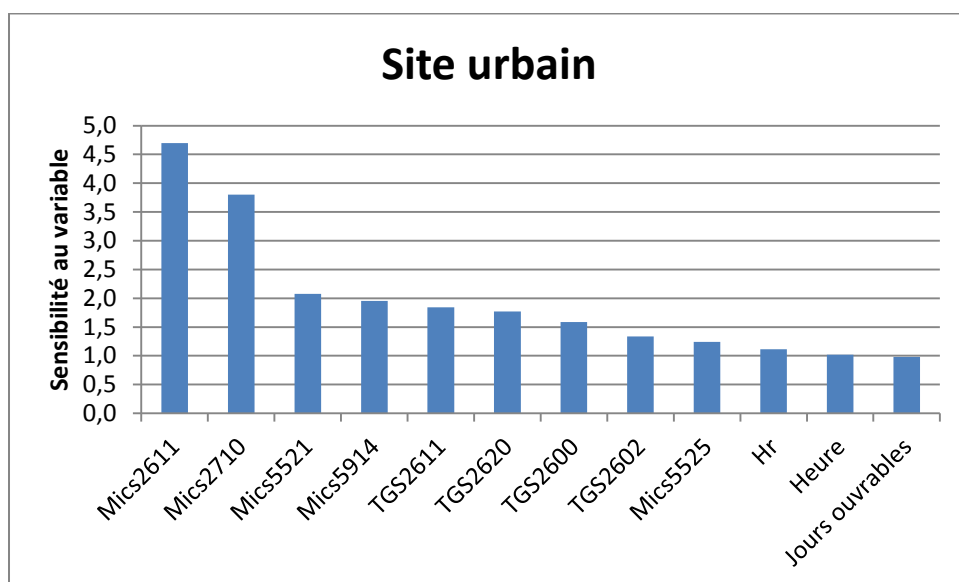


Figure 3-34 : Sensibilité du modèle aux variables, exemple d'un modèle calculé sur un site urbain.

Par contre, dans les stations trafic, les capteurs dédiés à la détection des COV comme le TGS2602 et le MICS5521 présentent des valeurs de sensibilités les plus élevées (3,7 et 2,2 respectivement) tandis que les variables de l'heure de jour et MICS2611 montrent les valeurs les plus basses (0,98 et 0,89). Ce qui implique que dans une station trafic, la variable des jours ouvrés apporte aux modèles plus d'information que le capteur d'O₃ MICS2611 (sans doute à cause de l'absence d'ozone) comme montre la figure 3-35.

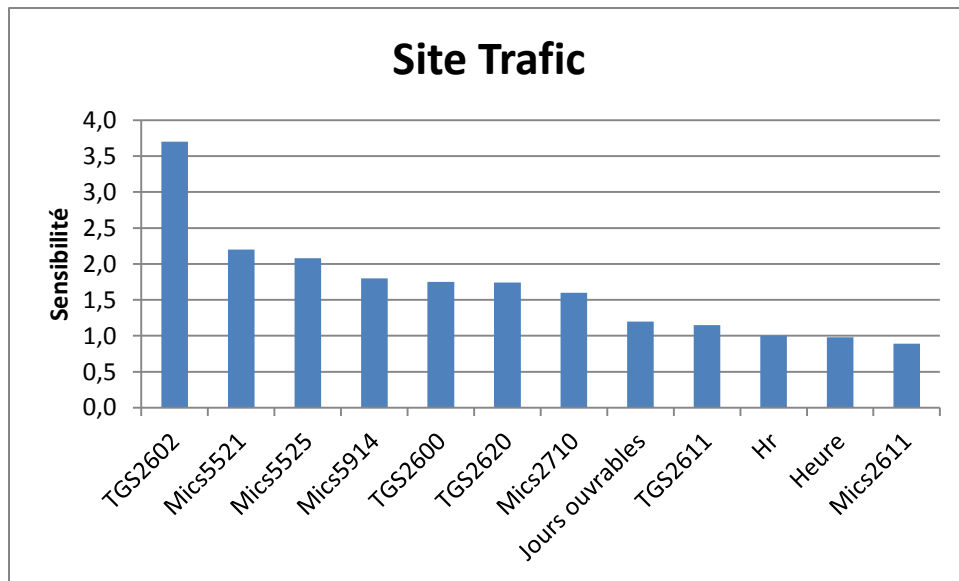


Figure 3-35 : Sensibilité du modèle aux variables, exemple d'un modèle calculé sur un site trafic.

Pour un modèle calculé sur plusieurs stations, toutes les variables possèdent des valeurs proches les unes des autres avec une différence de 0,05 entre la 1^{ère} et la dernière sensibilité. Ce qui indique l'importance de prendre en compte toutes les variables que nous avons décidé d'inclure comme données d'entrée pour obtenir un modèle universelle.

5.6. Validité des prédictions

Un des aspects négatifs d'une modélisation à base de réseaux de neurones, est la non capacité des modèles à détecter des événements inconnus, comme la présence d'un type inconnu de pollution, la présence d'interférent gazeux, ou la défaillance d'un capteur. De tels cas se traduiront toujours par des résultats a priori crédibles en phase de prédiction, et les anomalies ne seront donc pas détectées.

Une solution proposée est l'utilisation des résidus, ce qui est possible avec les méthodes de type PLS. Le résidu (en fait, la moyenne quadratique des résidus des X-variables) sur les variables d'entrées d'un modèle est le taux d'informations non utilisées durant une prédiction. Une valeur forte de résidu indique une quantité élevée d'informations non utilisée par le modèle au cours de la prédiction, ce qui est anormal. La suspicion de présence d'un

évènement anormal permet alors d'invalider la mesure. Mais les réseaux de neurones, choisis car plus performants que le PLS, ne permettent pas une telle démarche. Nous avons donc développé, en complément du réseau de neurone utilisé pour la prédiction, des modèles PLS utilisés uniquement pour l'examen des résidus, conjointement au modèle ANN.

Nous reprenons l'exemple montré dans la partie 3.2. de ce chapitre sur les données des capteurs (figure 3-12). Les points aberrants éliminés avant le traitement des données contiennent des informations sur la présence de polluants ou de pollutions inconnues au système (probablement passage d'un véhicule à proximité de la station de mesure), ces points seront un bon cas pour tester la non validité d'une prédiction.

Un modèle PLS a été construit sur les valeurs traitées et filtrées en retirant les points aberrants. Les résidus sont faibles sur les données moyennées, et majoritairement compris entre -5 et +5. Lorsque ce modèle a été appliqué aux données non moyennées, nous observons que ce résidu a augmenté de façon importante dépassant parfois la valeur de 40. La figure 3-36 montre les réponses brutes des capteurs en conductance. Ainsi que le carré du résidu calculé sur ces données brutes. Nous pouvons ainsi corréler les pics présents dans le résidu aux pics présents sur les réponses des capteurs. Ces pics proviennent d'évènements anormaux au voisinage de la station, comme le passage de véhicules ou encore les émissions de déchèteries voisines de la station de Saint Amand. Nous pouvons ainsi fixer un seuil de ce résidu au-delà duquel nous pouvons supposer que la prédiction risque d'être fausse (car on est en présence non survenu lors de l'apprentissage), et donc invalider la mesure.

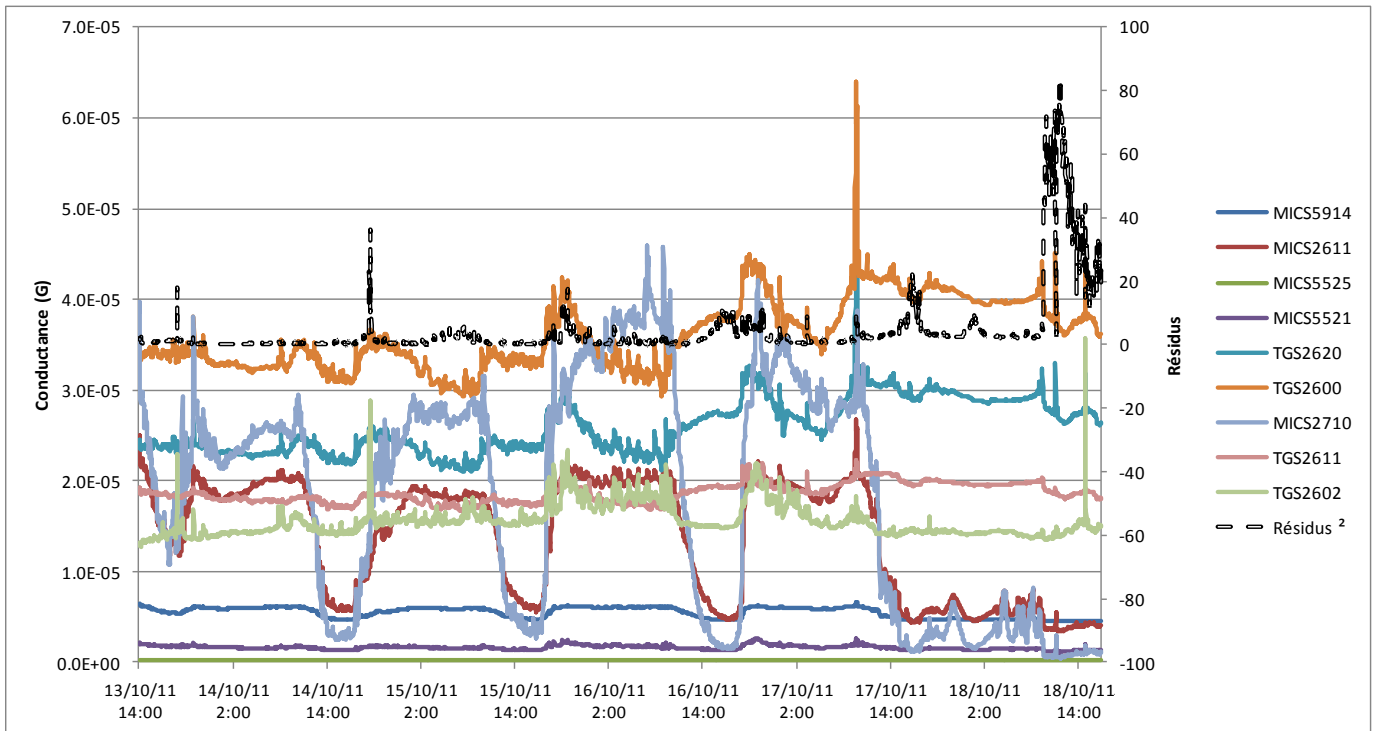


Figure 3-36 : variation de la conductance des capteurs (échelle de gauche) et résidu calculé (en noir, échelle de droite).

Nous avons aussi simulé une panne d'un capteur et étudié le comportement des résidus après cette panne. Nous avons choisi le capteur Mics2611, et avons forcé sa résistance à l'infini (simulant ainsi une rupture du capteur ou de sa résistance de chauffage), avec les mêmes données que précédemment. La figure 3-37 montre la variation des résidus calculés avant le début de la panne du capteur et après. Nous remarquons alors que les valeurs des résidus ont changé et passent largement et durablement au-delà de 5. Nous pouvons donc ici aussi introduire un seuil de résidus permettant d'invalidier la mesure.

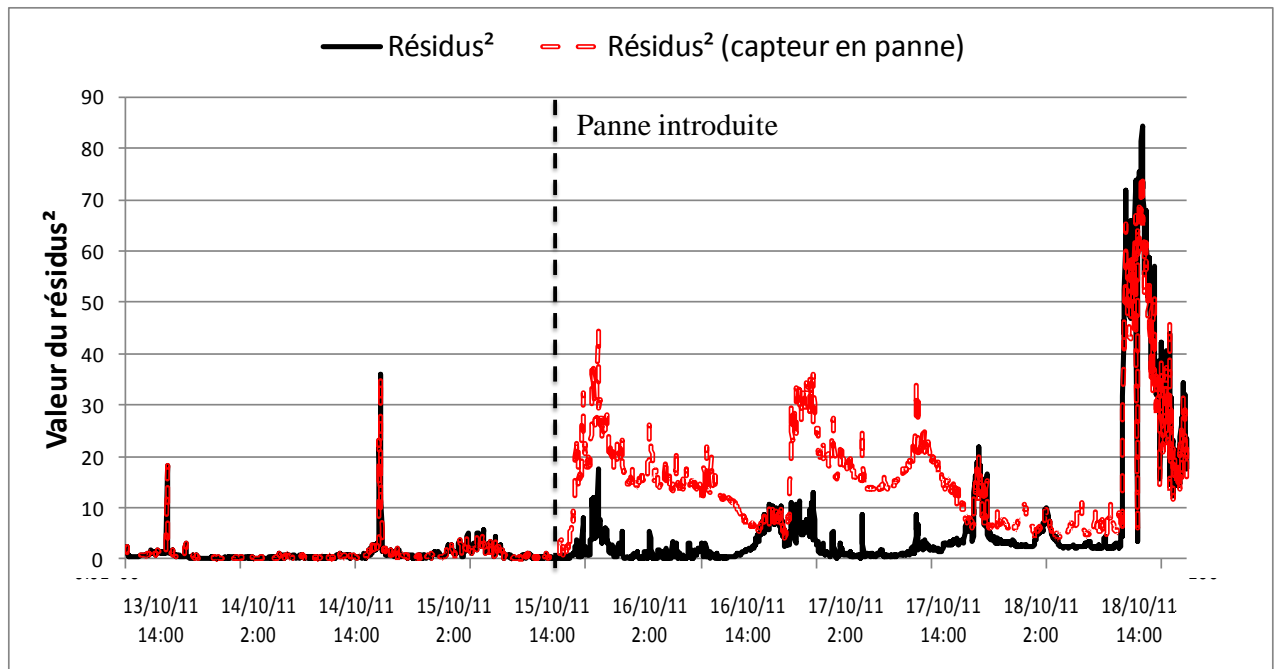


Figure 3-37 : Variation des résidus avant et après l'introduction de la panne.

6. Conclusion du chapitre

La détection de la pollution atmosphérique à l'aide des réseaux de capteurs a été étudiée dans ce chapitre. Une approche novatrice associant une analyse qualitative puis une analyse semi-quantitative a été adoptée. Pour atteindre, cet objectif, les concentrations des polluants dans l'air issues des stations de mesures des AASQA (Associations Agréées de Surveillance de la Qualité de l'Air) ont été étudiées et traitées afin de permettre une catégorisation adéquate des pollutions et en particulier afin de définir les seuils des concentrations de polluants définissant les diverses catégories de pollution. Les catégories définies sont « trafic », « urbain », et « photochimique », un indice de pollution a été également introduit en se basant sur l'indice ATMO, cet indice peut indiquer la sévérité de la pollution prépondérante. Ces catégories et cet indice ont été comparés avec les réponses des capteurs des réseaux des capteurs de gaz pour classer et identifier les pollutions de l'air.

Des modèles de comportement basés sur l'utilisation des réseaux de neurones ont été calculés à base, ils permettent de classer les différentes catégories de pollution en indiquant le

type et l'indice de la pollution prépondérante. Nous avons également étudié la possibilité d'établir des modèles universels pouvant soit être utilisés pour l'exploitation des données de différents réseaux de capteurs ou pour l'analyse des données obtenues sur différents sites sans apprentissage préalable spécifique au réseau de capteur ou au site d'étude. Les résultats obtenus ont alors mis en évidence une diminution des performances des modèles. Nous avons également mis en évidence qu'il est possible de capables de détecter des anomalies, soit de pollution, soit de comportement du réseau de capteurs, et alors d'invalidier la prédiction.

Remarquons enfin que les sites d'étude ayant permis de tester nos modélisations et prédictions étaient caractérisées par des niveaux de pollution relativement faibles, et que les seuils fixés et les modèles réalisés ne seront pas forcément adéquat pour des sites plus pollués.

En conclusion, il apparaît donc que nous sommes en mesure de réaliser des prédictions correctes même avec de faibles concentrations de polluants, ce qui n'était à priori pas évident au vu des gammes de concentrations spécifiées par les fabricants des capteurs.

Conclusion générale

L'objectif de ce travail était d'étudier la possibilité de développer des systèmes portables, *low-cost*, et autonomes pour la détection et la classification de différents types et niveaux de pollutions atmosphériques. Nous avons démontré que l'utilisation de capteurs de gaz à base de matériaux semi-conducteurs était adaptée à cette problématique. Neuf (9) capteurs commerciaux ont ainsi été sélectionnés, testés et étalonnés dans le laboratoire sous différents gaz présents naturellement dans un site pollué (CO, NO₂, CH₄, éthanol, toluène, H₂S, NH₃). Les capteurs utilisés ont présenté des réponses suffisamment intéressantes pour être exploitées pour l'application visée, bien que les concentrations des gaz générées sur nos bancs soient très basses en comparaison avec les gammes de détections annoncées par les constructeurs. Les inconvénients habituellement rencontrés avec ces outils de mesure, tels que la dérive des réponses au cours du temps, les problèmes de répétabilité d'un capteur à l'autre ou de reproductibilité d'un système à l'autre peuvent être soit corrigés, soit fortement atténués à l'aide de méthodes de normalisation et de standardisation adaptées. Comme cela a été démontré depuis la naissance des systèmes multi-capteurs, nous avons su tirer parti de la sélectivité partielle de ces capteurs pour la retourner à l'avantage des systèmes développés

grâce à l'analyse multi-variables. Cette approche a permis d'identifier des formes répétables dans les réponses des capteurs mis en œuvre dans les trois systèmes développés.

Jusqu'à présent, les travaux de recherche portant sur l'utilisation de réseaux des capteurs pour la détection d'espèces gazeuses avaient été menés en laboratoire, avec des générations de gaz synthétiques loin de la complexité d'une matrice d'espèces diverses et variées impliquées dans une pollution atmosphérique réelle. Nous avons donc décidé de confronter les systèmes conçus durant cette thèse à des situations de pollution atmosphérique réelle. Toutefois, la phase de mesures et tests en laboratoire s'est avérée indispensable pour la bonne compréhension du comportement des capteurs et définir une procédure d'utilisation optimale sur le terrain. Un développement technologique conséquent a été entrepris pour concevoir et réaliser trois cellules de mesure identiques avec capteurs et électroniques associées. Plusieurs campagnes de tests et de calibrage effectuées au laboratoire ont permis de valider le fonctionnement et la reproductibilité des cellules avant leur utilisation en sites réels de surveillance de la qualité de l'air.

La deuxième partie de ce travail de thèse a donc consisté à implémenter les trois systèmes multi-capteurs développés sur des sites sur lesquels la mesure d'un certain nombre de polluants réglementés est assuré par les AASQA afin d'enregistrer les réponses des capteurs sous différents types de pollution. Un historique d'un an de données des concentrations des polluants mesurés par les analyseurs des stations a été étudié pour chaque site. Cet historique a permis de définir 3 catégories de pollution (trafic, photochimique et urbain) à partir des formes journalières, hebdomadaires, et saisonnières de concentrations en polluants réglementés (NO , NO_2 et O_3). Ces catégories ont servi de base de travail pour définir un modèle capable de prédire la catégorie de la pollution à partir de la réponse seule des capteurs. Or l'utilisation d'une catégorisation provoque des problèmes de classification comme l'effet de seuil. Nous avons donc développé une nouvelle approche de type logique

floue dans laquelle nous traduirons les catégories de pollution en termes de degrés d'appartenances. Cette approche permet d'obtenir de bonnes performances de classification avec des taux de réussite de classification qui dépassent la plupart du temps 90%. Cependant cette méthode de prédiction reste purement catégorielle et n'apporte aucune information quantitative concernant le degré de pollution. Nous avons alors introduit une notion quantitative à la catégorisation des types de pollution en créant un Indice de Pollution (IP) s'inspirant de l'indice Atmo utilisé par les AASQA. Les nouveaux modèles ainsi développés ont montré un taux élevé de réussite (>90%) avec une faible erreur concernant la prédiction (faible RMSEP) de cet indice IP.

Nous avons par ailleurs montré que ces modèles de prédiction pouvaient avoir une portabilité inter-sites ou inter-réseau de capteurs, mais au prix de performances sensiblement moins bonnes. L'amélioration de la portabilité des systèmes reste dépendante des performances des capteurs, en particulier de leurs dérives dans le temps et des calibrages sous air « zéro » doivent être régulièrement effectués, ce qui est un peu en contradiction avec l'objectif initial de système portable et autonome. C'est un point important qui conditionne la finalisation d'éventuels transferts de technologies, et qui doit être pris en compte pour les futures recherches à mettre en place.

En ce qui concerne l'analyse des données, il ressort que l'utilisation des réseaux de neurones est préférable aux autres méthodes multi-variables car ils permettent de prendre en compte la non linéarité des capteurs et des catégories définies à l'aide de la logique floue. Par contre l'avantage d'autres méthodes comme les PCA et les PLS par rapport aux réseaux de neurones est d'avoir une information sur le résidu de prédiction et donc sur la fiabilité de la prédiction, nous avons alors montré que nous pouvons utiliser conjointement un réseau de neurones pour la classification / prédiction, et un examen des résidus issu de PLS afin de valider ou invalider cette prédiction.

En conclusion, nous avons montré dans ce travail la faisabilité d'un suivi certes non réglementaire mais néanmoins relativement performant de la pollution atmosphérique par des systèmes de mesures multi-capteurs à base des capteurs de gaz semi-conducteurs grâce à l'utilisation des méthodes multi-variables comme les réseaux de neurones. Nous avons déjà franchi une étape avec l'utilisation de capteurs commerciaux par rapport à des systèmes intégrant des capteurs de type « laboratoire ». Les futurs transferts de technologies réclament cependant encore une amélioration des performances des capteurs tant au niveau de leurs reproductibilités que de leurs dérives. Notre étude démontre que de tels systèmes multi-capteurs peuvent certainement être utilisés dans un avenir très proche pour le suivi de qualité de l'air extérieur ainsi que probablement de l'air intérieur. Ces systèmes pourraient aussi être utilisés pour l'identification de sources industrielles dans des sites « récepteurs » où cette identification est parfois délicate. L'utilisation de tels systèmes multi-capteurs intégrant des capteurs pourrait identifier des signatures spécifiques des différentes sources industrielles impactant potentiellement un site, ainsi de permettre l'attribution d'un indice de pollution à ces source.

Bibliographie

ADEME Classification et critères d'implantation des stations de surveillance de la qualité de l'air [Rapport]. - Paris : ADEME Editions, 2002.

Alexandre M. [et al.] Analysis of neural networks and analysis of feature selection with genetic algorithm to discriminate among pollutant gas. [Revue] // Sensors and Actuators B. - 2004. - Vol. 103. - pp. 122-128.

Althainz P. [et al.] Multisensor microsystem for contaminants in air [Revue] // Sensors and Actuators B 33. - 1996. - Vol. 33. - pp. 72-76.

Ampuero S. et Bosset J.O. The electronic nose applied to dairy products: a review [Revue] // Sensors and Actuators B. - 2003. - Vol. 94. - pp. 1-12.

Anderson H.R. Air pollution and mortality: A history [Revue] // Atmospheric Environment. - 2009. - Vol. 43. - p. 142.

Arquès P. La Pollution de l'air [Ouvrage]. - [s.l.] : Edisud, 1998.

Ash G. Les capteurs en instrumentation industrielle [Ouvrage]. - Paris : DUNOD, 2006.

Bitter F., Müller B. et Müller D. Estimation of odour intensity of indoor air pollutants from building materials with a multi-gas sensor system [Revue] // Building and Environment. - 2010. - Vol. 45. - pp. 197-204.

Brasan N. et Tomescu A. Calibration procedure for SnO₂ based gas sensors [Revue] // Thin Solid Films. - 1995. - Vol. 259. - pp. 91-95.

Breuil P. Analyses multivariées pour systèmes multi-capteurs [Section] // Les capteurs chimiques. - 1997.

Breuil P. Elaboration de couches minces de dioxyde d'étain sensible à l'action des gaz. Performances électriques et mécanismes réactionnels. (Thèse). - [s.l.] : IPN Grenoble et ENSM Saint Etienne, 1989.

Brezmes J. [et al.] Neural network based electronic nose for the classification of aromatic species [Revue] // Analytica Chimica ACTA . - 1997. - Vol. 348. - pp. 503-509.

Brudzewski K. [et al.] Chiral behavior of TGS gas sensors: Discrimination of the enantiomers by the electronic nose. [Revue] // Sensors and Actuators B . - 2007. - Vol. 122. - pp. 493-502.

Bruno L. Elaboration et caractérisation de couches minces de dioxyde d'étain obtenues par dépôt chimique en phase vapeur (CVD). Relation entre les propriétés structurales et électriques. Application à la détection de gaz. (thèse). - [s.l.] : IPN Grenoble et ENSM Saint Etienne, 1994.

C.E. COUNCIL DIRECTIVE of 15 July 1980 on air quality limit values and guide values for sulphur dioxide and suspended particulates. - [s.l.] : European Commission, 15 July 1980.

C.E. COUNCIL DIRECTIVE of 3 December 1982 on a limit value for lead in the air. - [s.l.] : European Commission, 3 December 1982.

C.E. COUNCIL DIRECTIVE of 7 March 1985 on air quality standards for nitrogen dioxide. - [s.l.] : European Commission, 7 March 1985.

C.E. Directive 1999/30/CE DU CONSEIL du 22 avril 1999 relative à la fixation de valeurs limites pour l'anhydride sulfureux, le dioxyde d'azote et les oxydes d'azote, les particules et le plomb dans l'air ambiant. - [s.l.] : Conseil Européen, 1999.

C.E. DIRECTIVE 2000/69/CE DU PARLEMENT EUROPEEN ET DU CONSEIL du 16 novembre 2000 concernant les valeurs limites pour le benzène et le monoxyde de carbone dans l'air ambiant.. - [s.l.] : Conseil Européen, 2000.

C.E. DIRECTIVE 2002/3/CE DU PARLEMENT EUROPEEN ET DU CONSEIL du 12 février 2002 relative à l'ozone dans l'air ambiant. - [s.l.] : Conseil Européen, 2002.

C.E. DIRECTIVE 2004/107/CE DU PARLEMENT EUROPEEN ET DU CONSEIL DE 15 décembre 2004 concernant l'arsenic, le cadmium, le mercure, le nickel et les hydrocarbures aromatiques polycycliques dans l'air ambiant. - [s.l.] : Conseil Européen, 2004. - pp. 23/3 - 23/16.

C.E. DIRECTIVE 2008/50/CE DU PARLEMENT EUROPÉEN ET DU CONSEIL du 21 mai 2008 concernant la qualité de l'air ambiant et un air pur pour l'Europe. - [s.l.] : Conseil Européen, 2008. - pp. 152/1 - 152/44.

C.E. DIRECTIVE 92/72/CEE DU CONSEIL du 21 septembre 1992 concernant la pollution de l'air par l'ozone. - [s.l.] : Conseil Européen, 1992. - pp. 297/1 - 297/7.

Cai J. et Levy D.C. Using stationary electronic noses network to locate dynamic odour source position [Conférence] // International conference on integration technology. - Shenzhen, China : IEEE, 2007. - pp. 793-798.

Capone S. [et al.] Analysis of vapours and foods by means of an electronic nose based on a sol-gel metal oxide sensors array [Revue] // Sensors and Actuators B . - 2000. - Vol. 69. - pp. 230-235.

Cevoli C. [et al.] Classification of Pecorino cheeses using electronic nose combined with artificial neural network and comparison with GC-MS analysis of volatile compounds. [Revue] // Food chemistry. - 2011. - Vol. 129. - pp. 1315-1319.

Chaiboun A. [et al.] A logarithmic multiparameter model using gas sensor main and cross sensitivities to estimate gas concentrations in a gas for SnO₂ gas sensors [Revue] // Sensors and Actuators B . - 2007. - Vol. 123. - pp. 1064-1070.

De Vito S. [et al.] CO, NO₂ and NO_x urban pollution monitoring with on field calibrated electronic nose by automatic bayesian regulation. [Revue] // Sensors and Actuators B. - 2009. - Vol. 143. - pp. 182-191.

Debliquy M. Capteurs chimiques [Revue] // Technique de l'ingénieur. - 2010. - R420.

Delpha C., Siadat M. et Lumbreras M. Humidity dependence of a TGS gas sensor array in an air-conditioned atmosphere [Revue] // Sensors and Actuators B. - 1999. - Vol. 59. - pp. 255-259.

Dimo I. [et al.] Studying the sensitivity of pollutants' concentrations caused by variations of chemical rates [Revue] // Journal of Computational and Applied Mathematics. - 2010. - Vol. 235. - pp. 391-402.

El Yamani M. Environnements et milieux [Rapport]. - 2006.

Flitti F. [et al.] Drift invariant Gas Recognition Technique For On Chip Tin Oxide Gas Sensor array [Conférence] // 4th IEEE International Symposium on Electronic Design, Test & Applications. - 2008.

FNORS Santé et environnement dans les régions de France [Rapport]. - [s.l.] : Ministère de la santé, de la Jeunesse, des Sports et de la vie associative, Septembre 2008.

Fort A. [et al.] Modeling of the influence of H₂O on metal oxide sensor response to CO. [Revue] // Sensors and Actuators B. - 2011. - Vol. 159. - pp. 82-91.

Fu J. [et al.] A pattern recognition method for electronic noses based on an olfactory neural network [Revue] // Sensors and Actuators B. - 2007. - Vol. 125. - pp. 489-497.

Garcia M. [et al.] Electronic nose for wine discrimination [Revue] // Sensors and Actuators B. - 2006. - Vol. 113. - pp. 911-916.

Gardner H., Shurmera V. et Julian W. Odour discrimination with an electronic nose [Revue] // Sensors and Actuators B: Chemical. - 1992. - Vol. 8. - pp. 1-11.

Gardner J.W. et Barlett P.N. Electronic noses, principles and applications. [Ouvrage]. - [s.l.] : Oxford university press, 1999.

Getino J. [et al.] Environmental applications of gas sensor arrays : combustion atmospheres and contaminated soils [Revue] // Sensors and Actuators B. - 1999. - Vol. 59. - pp. 249-254.

Getino J. [et al.] Environmental applications of gas sensor arrays: Combustion atmospheres and contaminated soils [Revue] // Sensors and Actuators B. - 1999. - Vol. 59. - pp. 249-254.

Haddi Z. [et al.] A portable electronic nose system for the identification of cannabis-based drugs [Revue] // Sensors and Actuators B. - 2011. - Vol. 155. - pp. 456-463.

- Hadfield L. et Seaton R.A.F.** A co-evolutionary model of change in environmental management. [Revue] // *Futures*. - 1999. - Vol. 31. - pp. 577-592.
- Hai Z. et Wang J.** Electronic nose and data analysis for detection of maize oil adulteration in sesame oil [Revue] // *Sensors and Actuators B*. - 2006. - Vol. 119. - pp. 449-455.
- Hartman J.D.** A possible objective method for rapid estimation of flavors in vegetables. [Revue] // *Proc. Am. Soc. Hort. Sci* 64. - 1954. - p. 335.
- Helli O., Siadat M. et Lumberas M.** Qualitative and quantitative identification of H₂S/NO₂ gaseous components in different reference atmospheres using metal oxide sensor array [Revue] // *Sensors and Actuators B*. - 2004. - Vol. 103. - pp. 403-408.
- Hines E.L., Llobet E. et Gardner J.W.** Electronic noses: a review of signal processing techniques [Revue] // *Circuits, Devices and Systems, IEE Proceedings*. - 1999. - Vol. 146. - pp. 297-310.
- Holmberg M. [et al.]** Identification of paper quality using a hybrid electronic nose [Revue] // *Sensors and Actuators B*. - 1995. - Vol. 26-27. - pp. 246-249.
- Houdret J.L., Mace T. et Jean P.** Réglementation et normalisation [Rapport]. - [s.l.] : LCSQA, 2008.
- Hunter L.J., Johnson G.T. et Watson I.D.** An investigation of three-dimensional characteristics of flow regimes within the urban canyon [Revue] // *Atmospheric Environment Part B : Urban Atmosphere*. - 1992. - Vol. 26 B. - pp. 425-432.
- Jaffrezic-Renault N., Martelet C. et Clechet P.** Capteurs chimiques et biochimiques [Revue] // *Techniques de l'ingénieur*. - 1994. - R420.
- Johnsons W.B. [et al.]** An urban diffusion simulation model of carbon monoxide [Revue] // *Journal of Air Pollution Controlling association*. - 1973. - Vol. 23. - pp. 490-498.
- Kamionak M., Breuil P. et Pijolat C.** Atmospheric pollution measurements with a multi-materials sensing device [Revue] // *Materials Science and Engineering C*. - 2006. - Vol. 26. - pp. 290-296.

Kamionka M. Développement de systèmes multicapteurs et multivariables pour la mesure en continue de polluants atmosphériques. (Thèse) // Thèse, ENSM Saint-Etienne. - Saint Etienne : [s.n.], 2005.

Kamionka M., Breuil P. et Pijolat C. Calibration of a multivariate gas sensing device for atmospheric pollution measurement. [Revue] // Sensors and Actuators B. - 2006. - Vol. 118. - pp. 323-327.

Ködderitzsch P. [et al.] Sensor array based measurement technique for fast-responding cigarette smoke analysis [Revue] // Sensors and Actuators B. - 2005. - Vol. 107. - pp. 479-489.

Koga M. [et al.] Determination of ppt levels of atmospheric volatile organic compounds in Yakushima, a remote south-west island of Japan [Revue] // Microchemical Journal. - 2001. - Vol. 68. - pp. 257-264.

Korotcenkov G. et Cho B.K. Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey) [Revue] // Sensors and Actuators B. - 2011. - Vol. 156. - pp. 527-538.

Kulkarni N. et Grigg J. Effect of air pollution on children [Conférence] // Symposium: social paediatrics. - [s.l.] : Elsevier, 2008. - pp. 238-243.

Kuske M. [et al.] Fuzzy k-NN applied to moulds detection [Revue] // Sensors and Actuators B. - 2005. - Vol. 106. - pp. 52-60.

Lalauze R. Interaction gaz-solide [Ouvrage]. - [s.l.] : CMC2, 1997. - pp. 28-38.

Le Cloirec P. COV (Composés organiques volatils) [Revue] // Techniques de l'ingénieur. - 2004. - G 1 835.

Lee D. [et al.] Recognition of volatile organic compounds using SnO₂ sensor array and pattern recognition analysis [Revue] // Sensors and Actuators B. - 2001. - Vol. 77. - pp. 228-236.

Liu W.T. [et al.] Diagnosis of air quality through observation and modelling of volatile organic compounds (VOCs) as pollution tracers [Revue] // Atmospheric Environment. - 2012. - Vol. 55. - pp. 56-63.

Llop S. et Ballester F. Ambient air pollution and annoyance responses from pregnant women [Revue] // Atmospheric Environment. - 2008. - Vol. 42. - pp. 2982-2992.

Lozano J. [et al.] Electronic nose for wine ageing detection [Revue] // Sensors and Actuators B. - 2008. - Vol. 133. - pp. 180-186.

Ma Y. [et al.] Air Pollution Monitoring and Mining Based on Sensor Grid in London [Revue] // Sensors. - 2008. - pp. 3601-3623.

Martin M., Santos J. et Agapito J. Application of artificial neural networks to calculate the partial gas concentrations in a mixture [Revue] // Sensors and Actuators B 77. - 2001. - pp. 468-471.

Mathé F. et Houdret J.L. Surveillance de la qualité de l'air: le dispositif français [Revue] // Techniques de l'ingénieur. - 2009. - G 1 632.

Mathé F. Evolution de la classification et des critères d'implantation des stations de mesure de la qualité de l'air - Participation à la reactualisation du guide de classification des stations [Rapport]. - [s.l.] : Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA), 2010.

Mayer H. Air pollution in cities [Revue] // Atmospheric Environment. - 1999. - Vol. 33. - pp. 4029-4037.

Menil F. Capteurs à électrolyte solide [Ouvrage]. - [s.l.] : Lavoisier, 2008.

Mielle P. et Marquis F. One-sensor electronic olfactometer for rapid sorting of fresh fruit juices [Revue] // Sensors and Actuators B. - 2001. - Vol. 76. - pp. 470-476.

Ministère du développement durable La réglementation en matière de qualité de l'air [En ligne]. - 12 janvier 2010. - <http://www.developpement-durable.gouv.fr/La-reglementation-en-matiere-de.html>.

Monk P.S. [et al.] Atmosphéric composition change: global and regional air quality [Revue] // Atmospheric environment Vol.43. - 2009.

Montméat P. Role d'éléments métalliques sur les mécanismes de détection d'un capteur de gaz à base de dioxyde d'étain. Application à l'amélioration de la sélectivité à l'aide d'une membrane de platine. - [s.l.] : ENSM Saint-Etienne, 2002.

Morvan M. [et al.] Electronic-nose systems for control quality applications in automotive industry [Revue] // Sensors and Actuators B 69. - 2000. - pp. 384-388.

Munz P. [et al.] WHEN ZOMBIES ATTACK!: MATHEMATICAL MODELLING OF AN OUTBREAK OF ZOMBIE INFECTION [Revue] // Infectious Disease Modelling Research Progress. - 2009. - pp. 133-150.

Nakata S., Ozaki E. et Ojima N. Gas sensing based on the dynamic nonlinear responses of a semiconductor gas sensor: dependence on the range and frequency of a cyclic temperature change [Revue] // Analytica Chimica Acta 361. - 1998. - pp. 93-100.

Nake A. [et al.] Outdoor in situ monitoring of volatile emissions from wastewater treatment plants with two portable technologies of electronic noses [Revue] // Sensors and Actuators B 106. - 2005. - Vol. 106. - pp. 36-39.

Natale C. Di, Davide F. et A.D'Amico Patter recognition in gas sensing: well-stated techniques and advances [Revue] // Sensors and Actuators B 23. - 1995. - pp. 111-118.

Negri R.M. et Reich S. Identification of pollutant gases and its concentrations with a multisensor array [Revue] // Sensors and Actuators B. - 2001. - Vol. 75. - pp. 172-178.

Ngo K.A. Etude d'un système multicapteur pour la détection sélective des gaz. (thèse). - Marseille : Université Cezanne, 2006.

Nicolas J., Craffe F. et Romain A.C. Estimation of odor emission rate from landfill areas using the sniffing team method [Revue] // Waste Management 26. - 2006. - pp. 1259-1269.

OECD Environmental outlook to 2030 [Rapport]. - [s.l.] : OECD, 2008.

ORS Les composés organiques volatils (COV) [Rapport]. - Ile-de-France : Observatoire régional de santé d'Ile-de-France, décembre 2007.

Otto M. Chemometrics, Statistics and computer application in Analytical Chemistry [Ouvrage]. - [s.l.] : Wiley-VCH verlag GmbH & Co. KGaA, 2007.

Padilla M. [et al.] Drift compensation of gas sensor array data by Orthogonal Signal Correction [Revue] // Chemometrics and Intelligent Laboratory Systems 100. - 2010. - pp. 28-35.

Pascal L. Effet à court terme de la pollution atmosphérique sur la mortalité. [Revue] // Revue française d'allergologie Vol. 49. - 2009.

Peden D.B. Effect of air pollution in asthma and respiratory allergy [Revue] // Otolaryngology - Head and Neck Surgery . - 1996. - Vol. 114. - pp. 242-247.

Percy K.E. et Feretti M. Air pollution and forest health: toward new monitoring concepts [Revue] // Environmental Pollution Vol.30. - 2004. - pp. 113-126.

Perdreau N. Application des methodes d'analyse multivariées à la detection quantitative de gaz par microcapteurs à base de dioxyde d'étain. (thèse). - IPN Grenoble et EMSE Saint-Etienne : [s.n.], 2000.

Permuy A. [et al.] Capteurs à semi-conducteurs [Revue] // Techniques de l'ingénieur. - 1994. - Vol. 420. - p. 360.

Pijolat C. Etude des propriétés physico-chimiques et des propriétés électriques du dioxyde d'étain en fonction de l'atmosphère gazeuse environnante. Application à la détection sélective des gaz. (Thèse). - Grenoble : [s.n.], 1986.

Pijolat C. Micro-capteur de gaz élaboré a partir de matériaux solides [Section] // LES CAPTEURS CHIMIQUES. - [s.l.] : CMC2, 1997.

Principe J.C. [et al.] Neural and Adaptive Systems. Fundamentals Through Simulations. [Ouvrage]. - [s.l.] : JOHN WILE & SONS, INC., 2000.

Ramanathan V. et Feng Y. Air pollution, greenhouse gases and climate change: Global and regional perspectives [Revue] // Atmospheric Environment 43. - 2009. - pp. 37 - 50.

Rivière B. Optimisation du pricecde de serigraphie pour la réalisation de capteurs de gaz en couche épaisse. Etude de la compatibilité avec la technologie microélectronique. (Thèse). - [s.l.] : ENSM Saint-Etienne, 2004.

Romain A.C. [et al.] Monitoring the exhaust air of a compost pile with an e-nose and comparison with GC-MS data [Revue] // Sensors and Actuators B. - 2005. - Vol. 106. - pp. 317-324.

Romain A.C. [et al.] Use of a simple tin oxide sensor array to identify five malodours collected in the field [Revue] // Sensors and Actuators B. - 2000. - Vol. 62. - pp. 73-79.

Romain A.C. [et al.] Using the classification model of an electronic nose to assign unknown malodours to environmental sources and to monitor them continuously [Revue] // Sensors and Actuators B. - 2000. - Vol. 69. - pp. 366-371.

Romain A.C. et Nicolas J. Long term stability of metal oxide-based gas sensors for e-nose environmental applications: An overview. [Revue] // Sensors and Actuators B. - 2010. - Vol. 146. - pp. 502-506.

Saeed S., Abbas Z. et Gopal B. Experimental Use of Electronic Nose for Analysis of Volatile Organic Compound (VOC) [Revue] // IEEE. - 2009.

Sarry F. et Lumbreras M. Gas composition determination in an air conditioned system using a sensor array: characterization of three different TGS sensor. [Revue] // Sensors and Actuators B. - 1999. - Vol. 59. - pp. 94-99.

Sauvan M. Modification de la sélectivité de couches minces de dioxyde d'étain par l'ajout de couches superficielles, en vue de la réalisation de microcapteurs de gaz. (Thèse). - [s.l.] : IPN Grenoble et ENSM Saint-Etienne, 1999.

Schweizer-Berberich M. [et al.] Strategies to avoid VOC cross-sensitivity of SnO₂-based CO sensors [Revue] // Sensors and Actuators B. - 1998. - Vol. 58. - pp. 318-324.

Shaffer R. [et al.] A comparison study of chemical sensor array pattern recognition algorithms [Revue] // Analytica Chimica Acta. - 1999. - Vol. 384. - pp. 305-317.

Shurmer H.V., Corcoran P. et James M.K. Sensitivity enhancement for gas sensing and electronic nose applications [Revue] // Sensors and Actuators B: Chemical. - 1993. - Vol. 16. - pp. 256-259.

Sivacoumar R. [et al.] Air pollution modeling for an industrial complex and model performance evaluation [Revue] // Environmental pollution. - 2001. - Vol. 111. - pp. 471-477.

Skoog D.A. et Leary J.J. Principles of Instrumental Analysis [Ouvrage]. - International Edition : [s.n.], 1992.

Smith L. A tutorial on Principal Components Analysis [Ouvrage]. - 2002.

Sommer V. [et al.] Neural networks and abductive networks for chemical sensor signals: a case comparison [Revue] // Sensors and Actuators B. - 1995. - Vol. 28. - pp. 217-222.

Strobel P. [et al.] A portative gas recognition system based on metal oxide gas sensors array [Revue] // Sensors. - [s.l.] : Proceedings of IEEE, 2004. - Vol. 1. - pp. 123-126.

Szcsurek A. [et al.] Method of gas mixtures discrimination based on sensor array, temporal response and data driven approach [Revue] // Talanta. - 2011. - Vol. 83. - pp. 916-923.

Szczurek A. [et al.] A concept of a sensor system for determining composition of organic solvents [Revue] // sensors journal. - [s.l.] : IEEE , 2010. - 5 : Vol. 10. - pp. 924-933.

Szecowka P.M. [et al.] On reliability of neural network sensitivity analysis applied for sensors array optimisation [Revue] // Sensors and Actuators B. - 2011. - Vol. 157. - pp. 298-303.

Thery-Merland F. [et al.] Mise au point d'un multicapteur de gaz [Revue] // Techniques de l'ingénieur, IN 33-7, 6. - [s.l.] : Techniques de l'ingénieur, 2006. - IN 33-7, 6.

Tomic O. [et al.] Recalibration of a gas-sensor array system related to sensor replacement [Revue] // Analytica Chimica Acta. - 2004. - Vol. 512. - pp. 199-206.

Tomic O. [et al.] Standardization methods for handling instrument related signal shift in gas-sensor array measurement data [Revue] // Analytica Chimica Acta. - 2002. - Vol. 472. - pp. 99-111.

Toupance G. [et al.] Pollution atmosphérique gazeuse; Mesure des gaz [Revue] // Techniques de l'Ingénieur. - 2004. - P4031.

Toupance G. et Person A. Pollution atmosphérique gazeuse, mesure des gaz. [Revue] // Techniques de l'ingénieur. - 2004. - P 4 031.

Trincavelli M. [et al.] Towards Environmental Monitoring with Mobile Robots [Conférence] // IEEE/RSJ International Conference on Intelligent Robots and Systems. - Nice - France : [s.n.], 2008. - pp. 2210-2215.

Tsujita W. [et al.] Gas sensor network for air-pollution monitoring [Revue] // Sensors and Actuators B. - 2005. - Vol. 110. - pp. 304-311.

Valleron A. Etude et optimisation de capteurs de gaz à base de dioxyde d'étain (SnO₂) en conditions d'une ligne d'échappement automobile (Thèse). - EMSE Saint-Etienne : [s.n.], 2011.

Wilken W.F. et Hartman J.D. An electronic analogue for the olfactory process. [Revue] // Ann. NY Acad. Sci. - 1964. - Vol. 116. - pp. 608-612.

World S. Chemometrics: what do we mean with it and what do we want from it [Revue] // Chemometrics and intelligent Laboratory systems. - 1995. - Vol. 30. - pp. 109-115.

Yamazoe N. et Shimanoe K. New perspectives of gas sensor technology [Revue] // Sensors and Actuators B. - 2009. - Vol. 138. - pp. 100-107.

Yang B. [et al.] Quantification of H₂S and NO₂ using gas sensor arrays and an artificial neural network [Revue] // Sensors and Actuators B. - 2007. - Vol. 43. - pp. 235-238.

Zadeh L.A. Fuzzy sets [Revue] // Information and Control. - 1968. - 3 : Vol. 8. - pp. 338-353.

Zampolli S. [et al.] An electronic nose based on solid state sensor arrays for low-cost indoor air quality monitoring applications [Revue] // Sensors and Actuators B. - 2004. - Vol. 101. - pp. 39-46.

Publications et communications

Communications :

- The 14th international meeting on Chemical Sensors (20-22 mai 2012):

Titre: Development of a standardized multi-sensors system for on-line atmospheric pollution monitoring.

Auteurs: Zaher AL BARAKEH, Philippe BREUIL, Nathalie REDON, Christophe PIJOLAT, Nadine LOCOGE, Jean-Paul VIRICELLE

- 20th International Conference on Modelling, Monitoring and Management of Air Pollution (16-18 mai 2012):

Titre: A multi-sensors system for atmospheric pollution detection, a classification and indexation approach.

Auteurs: Zaher AL BARAKEH, Philippe BREUIL, Nathalie REDON, Christophe PIJOLAT, Nadine LOCOGE, Jean-Paul VIRICELLE

Publications :

- Sensors and Actuators B : Chemical (*en cours de publication*)

Titre: Development of a standardized multi-sensors system for on-line atmospheric pollution monitoring.

Auteurs: Zaher AL BARAKEH, Philippe BREUIL, Nathalie REDON, Christophe PIJOLAT, Nadine LOCOGE, Jean-Paul VIRICELLE

- Environmental Science and Technology (*en cours de publication*)

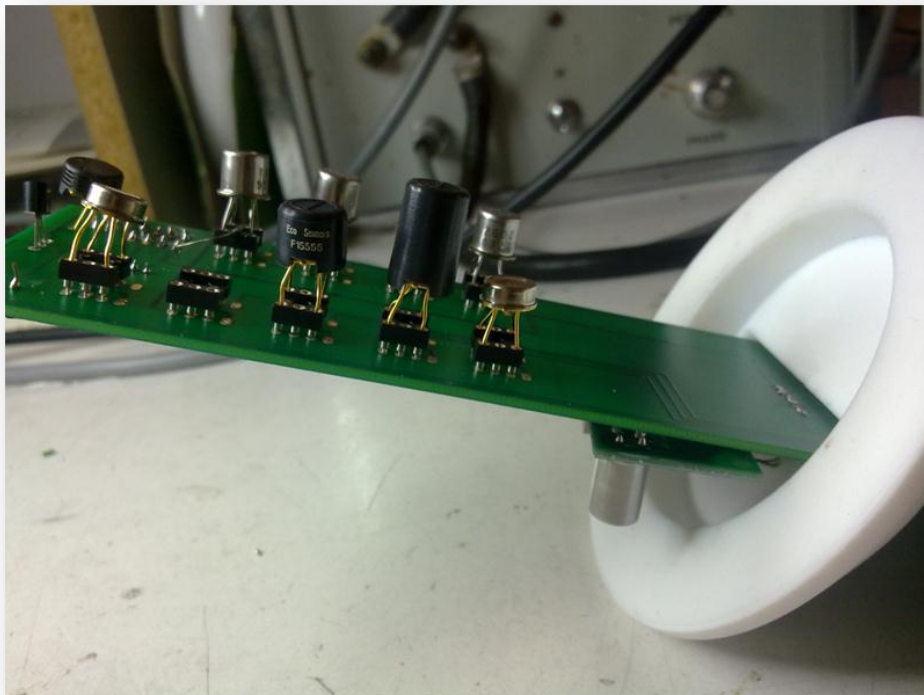
Titre: A multi-sensors system for atmospheric pollution detection, a classification and indexation approach.

Auteurs: Zaher AL BARAKEH, Philippe BREUIL, Nathalie REDON, Christophe PIJOLAT, Nadine LOCOGE, Jean-Paul VIRICELLE

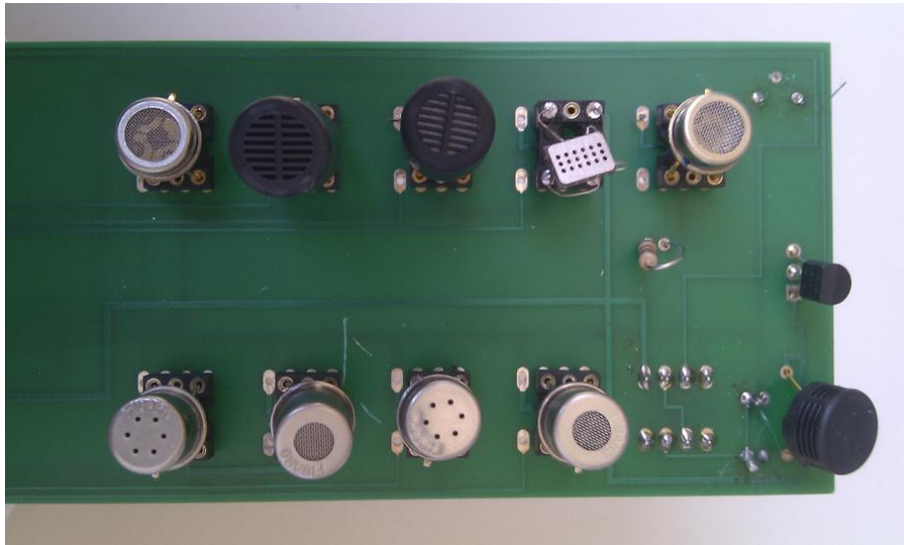
Annexes :

Annexe A : Systèmes multi-capteurs développés :

Vue profile du système développé :



Vue d'en haut du système développé :



Trois systèmes multi-capteurs branché sur le banc de test :



Système d'acquisition de données à utilisation terrain :



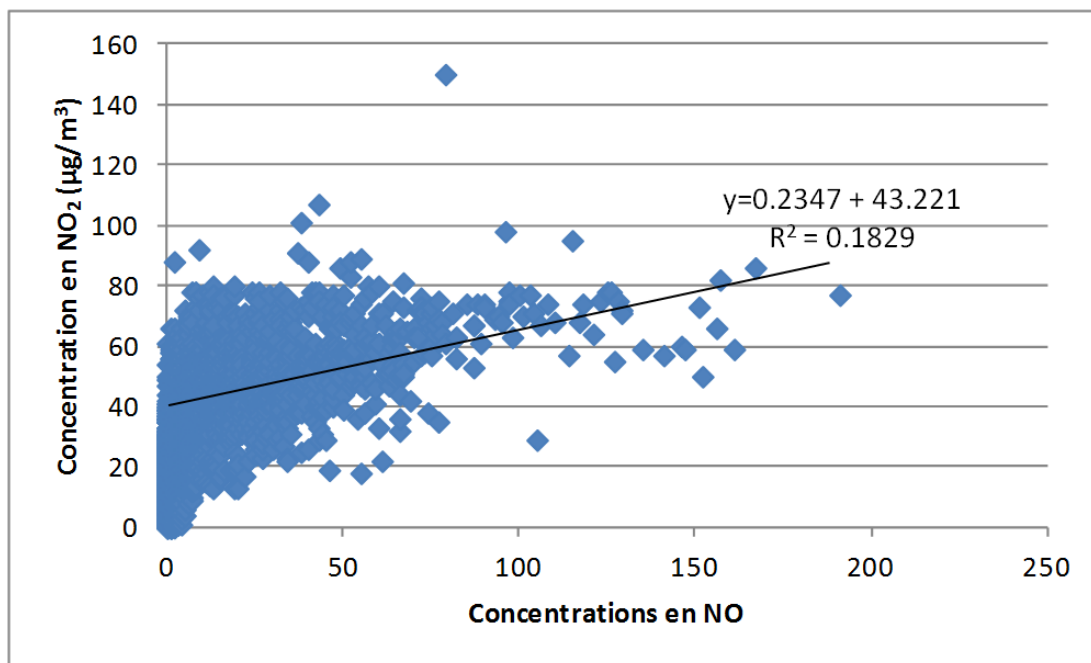
Photo d'une station de mesure de la qualité d'air :



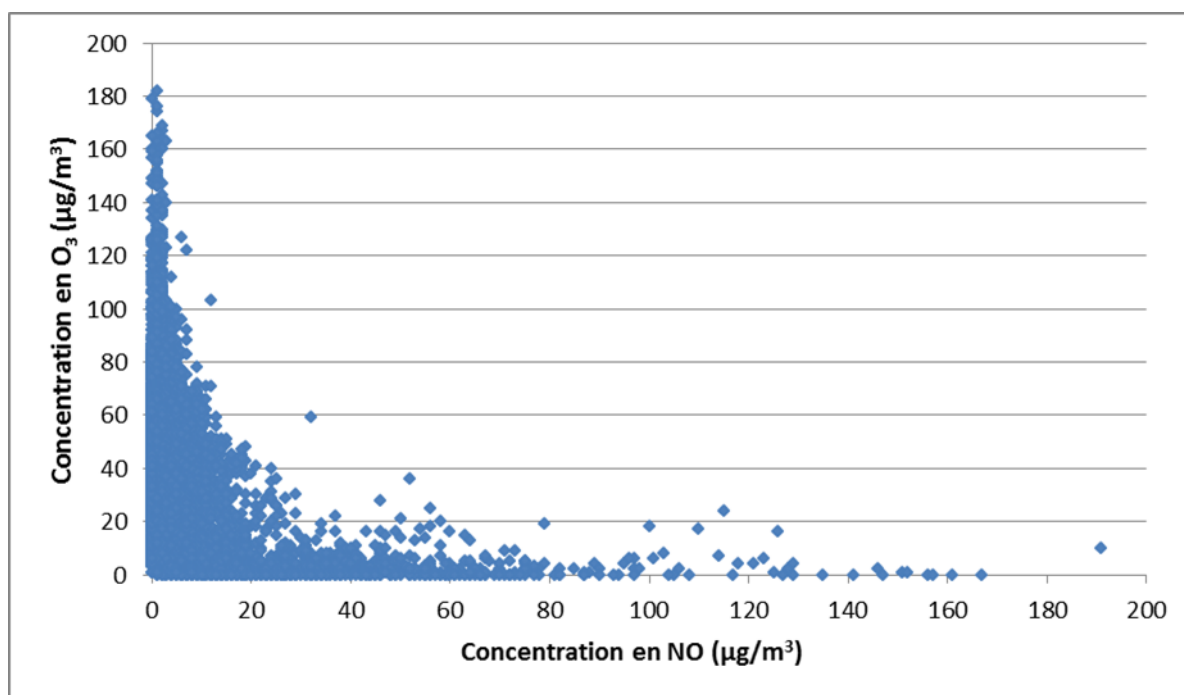
Annexe B : Corrélations entre les concentrations des polluants :

Station de Béthune (urbaine) :

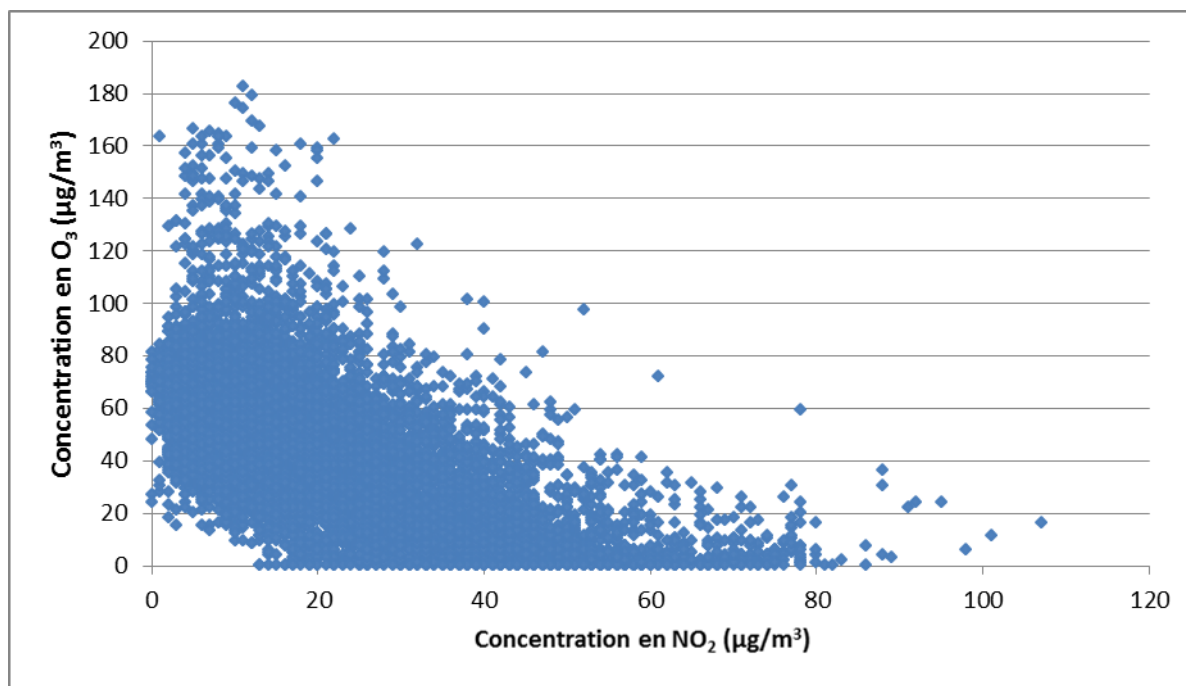
- Relation entre NO_2 et NO , y et R^2 établies pour $\text{NO} > 20 \mu\text{g}/\text{m}^3$:



- Relation entre O_3 et NO :

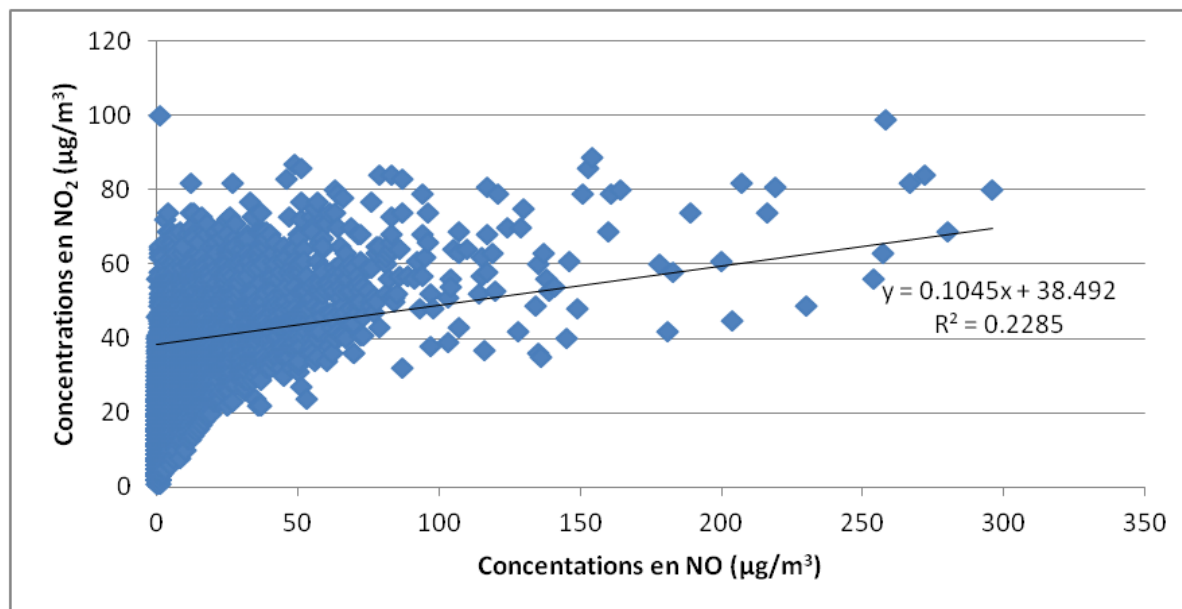


- Relation entre NO_2 et NO :

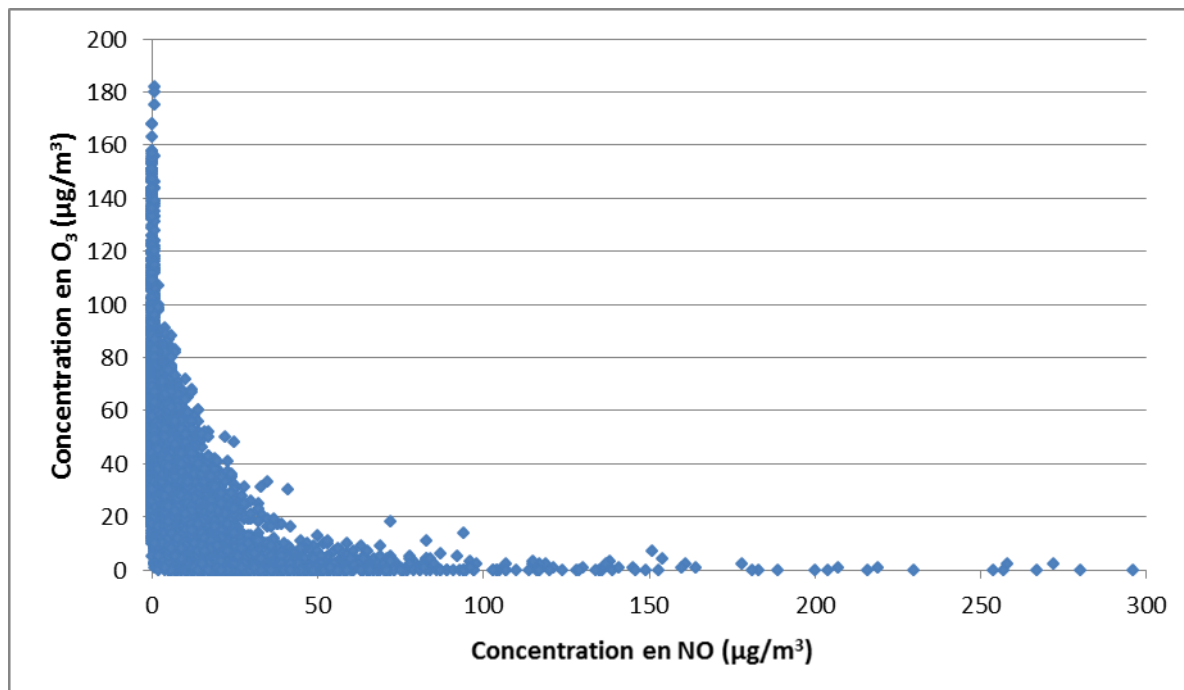


Station de Douai (urbaine) :

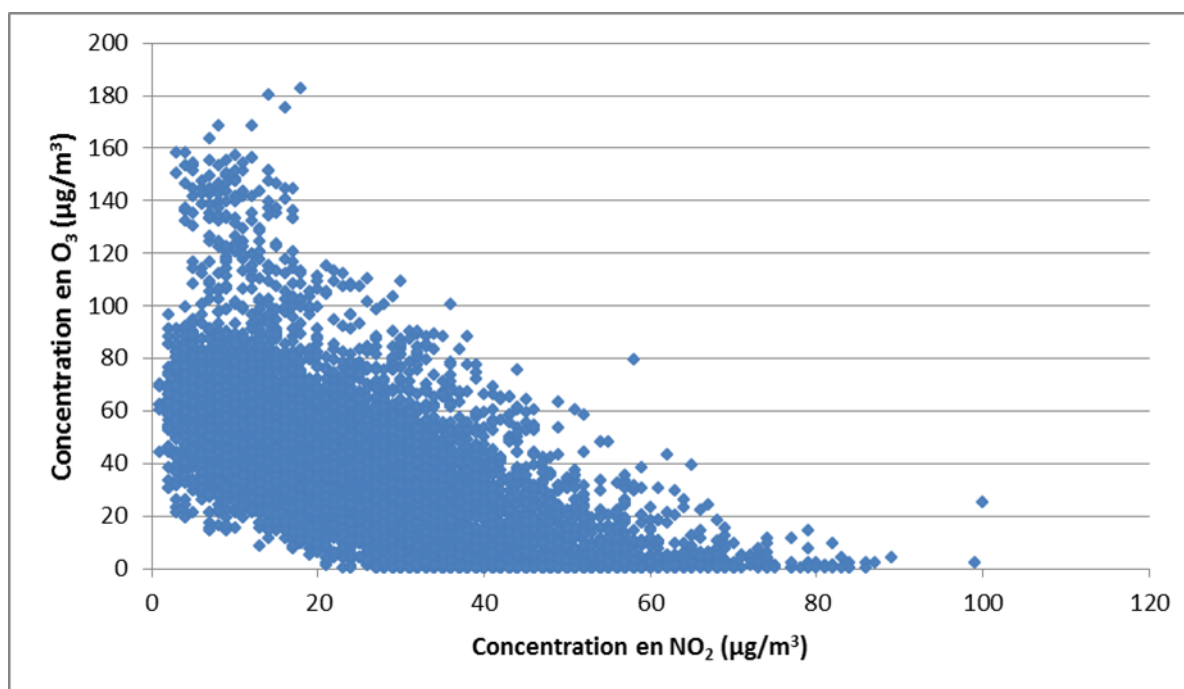
- Relation entre NO_2 et NO , y et R^2 établies pour $\text{NO} > 20 \mu\text{g/m}^3$:



- Relation entre O_3 et NO :

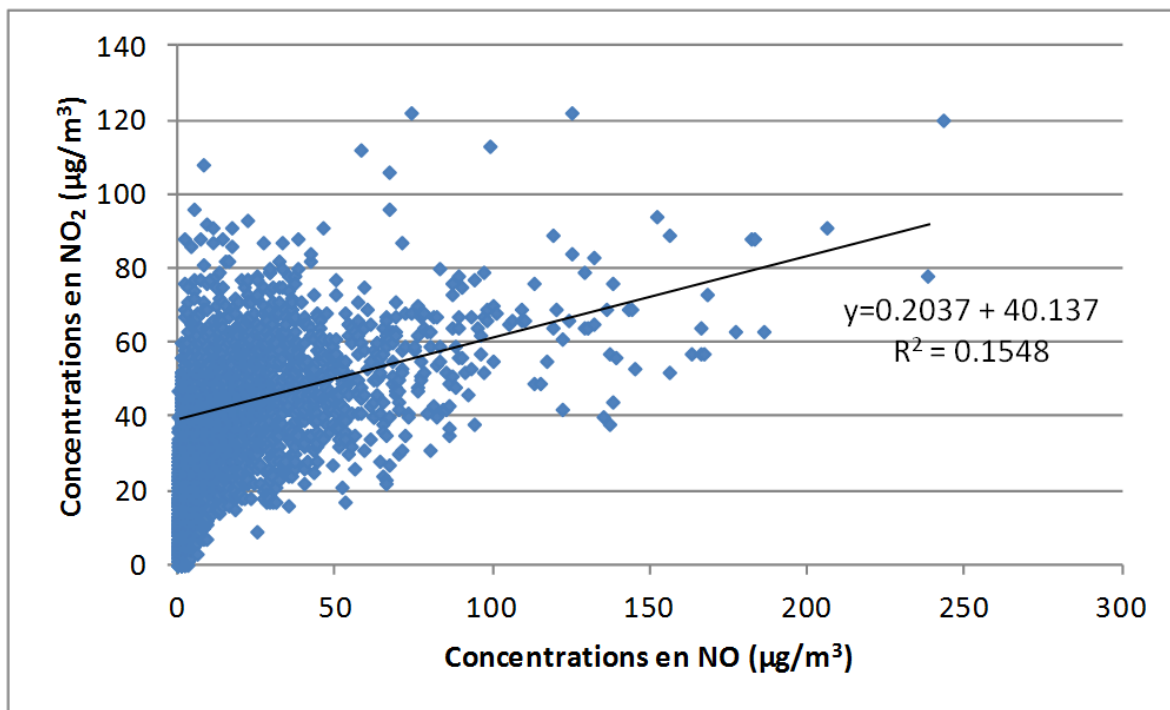


- Relation entre O_3 et NO_2 :

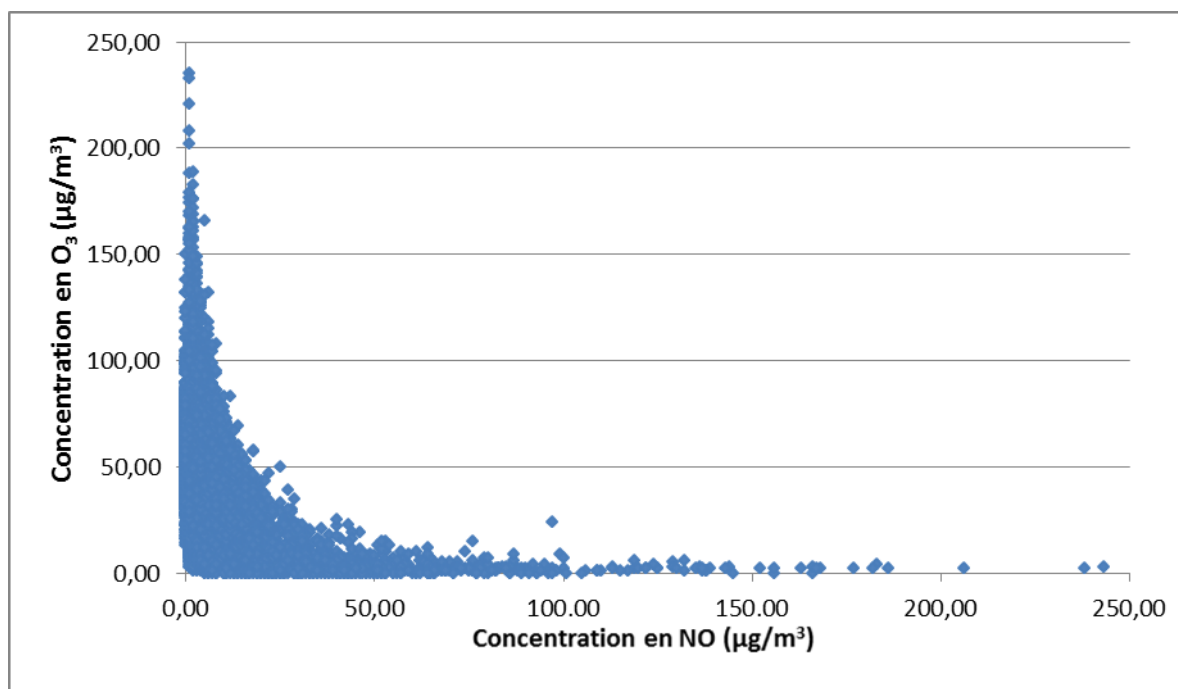


Station de Creil (urbaine) :

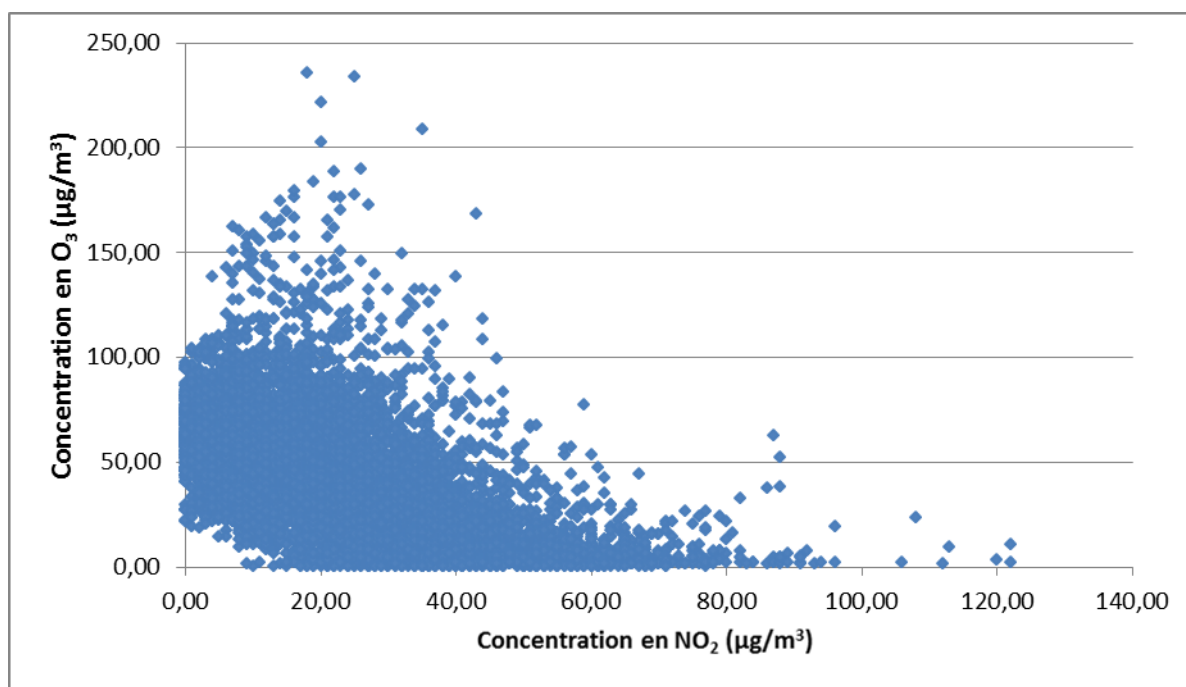
- Relation entre NO_2 et NO , y et R^2 établies pour $\text{NO} > 20 \mu\text{g}/\text{m}^3$:



- Relation entre O_3 et NO :

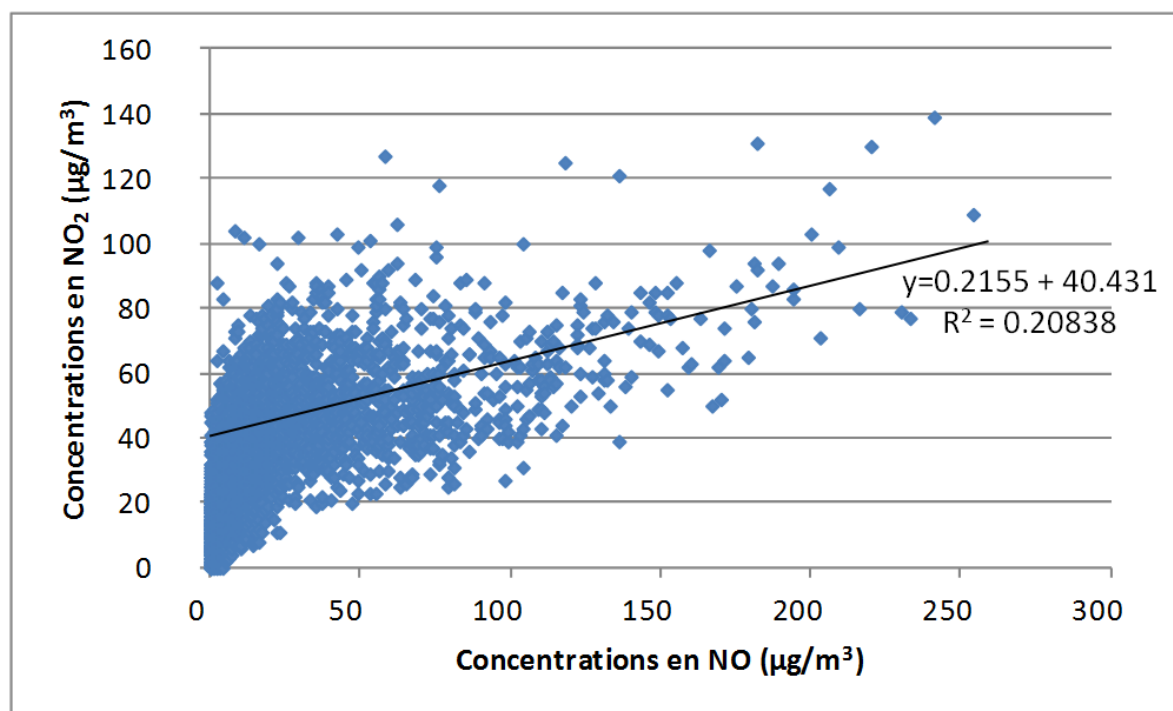


- Relation entre O_3 et NO_2 :

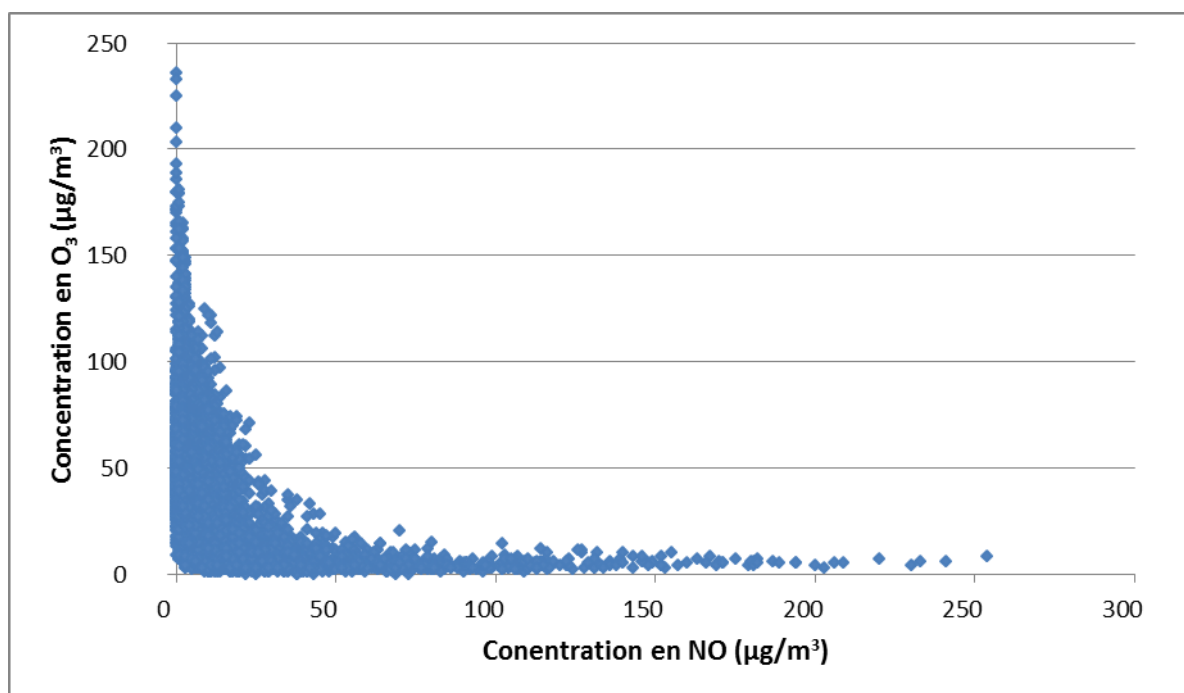


Station de Nogent sur Oise (périurbaine) :

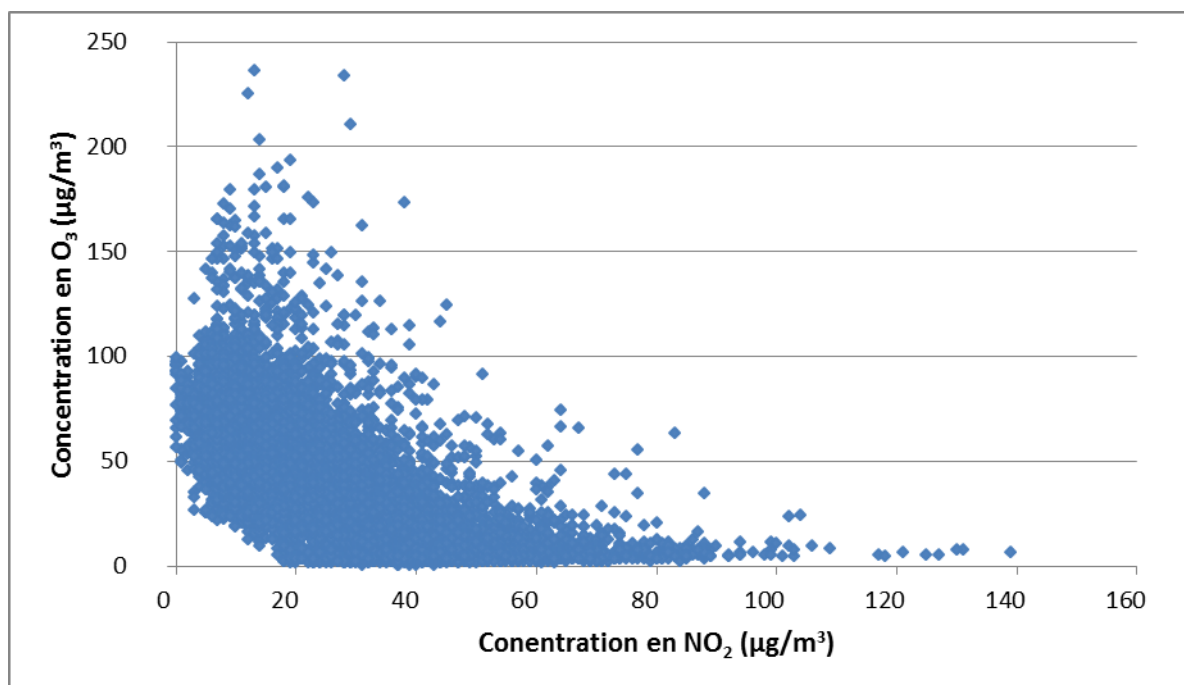
- Relation entre NO_2 et NO , y et R^2 établies pour $NO > 20 \mu g/m$:



- Relation entre O_3 et NO :

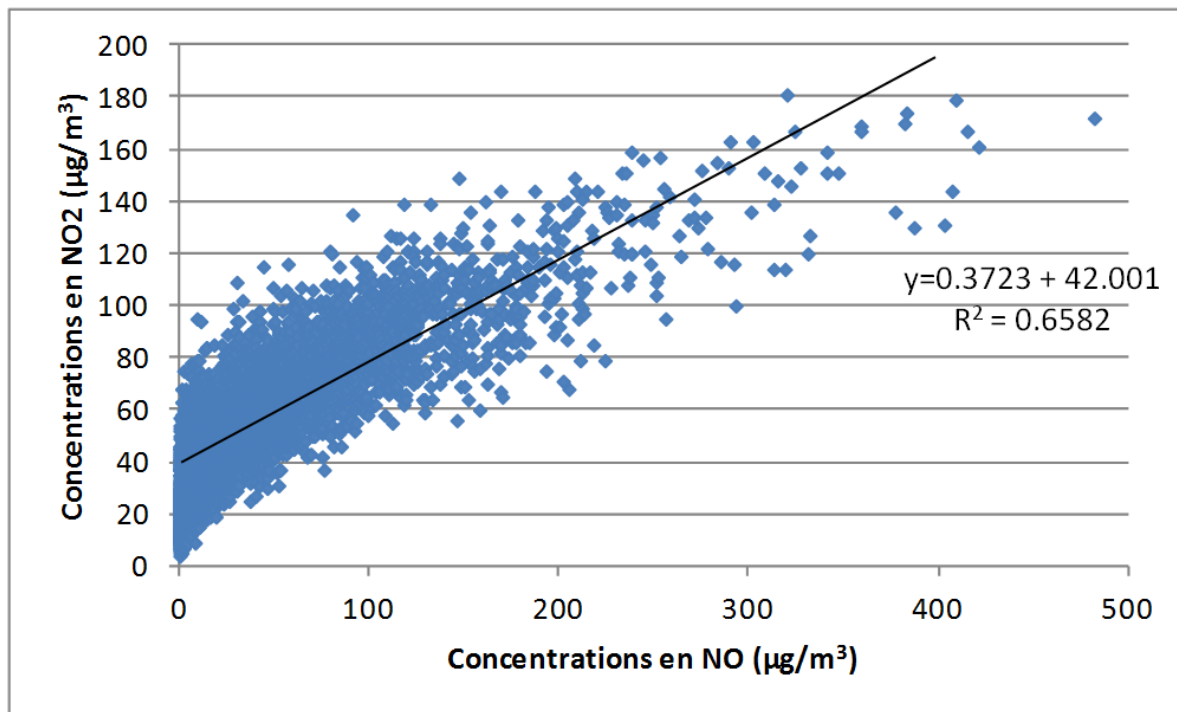


- Relation entre O_3 et NO_2 :



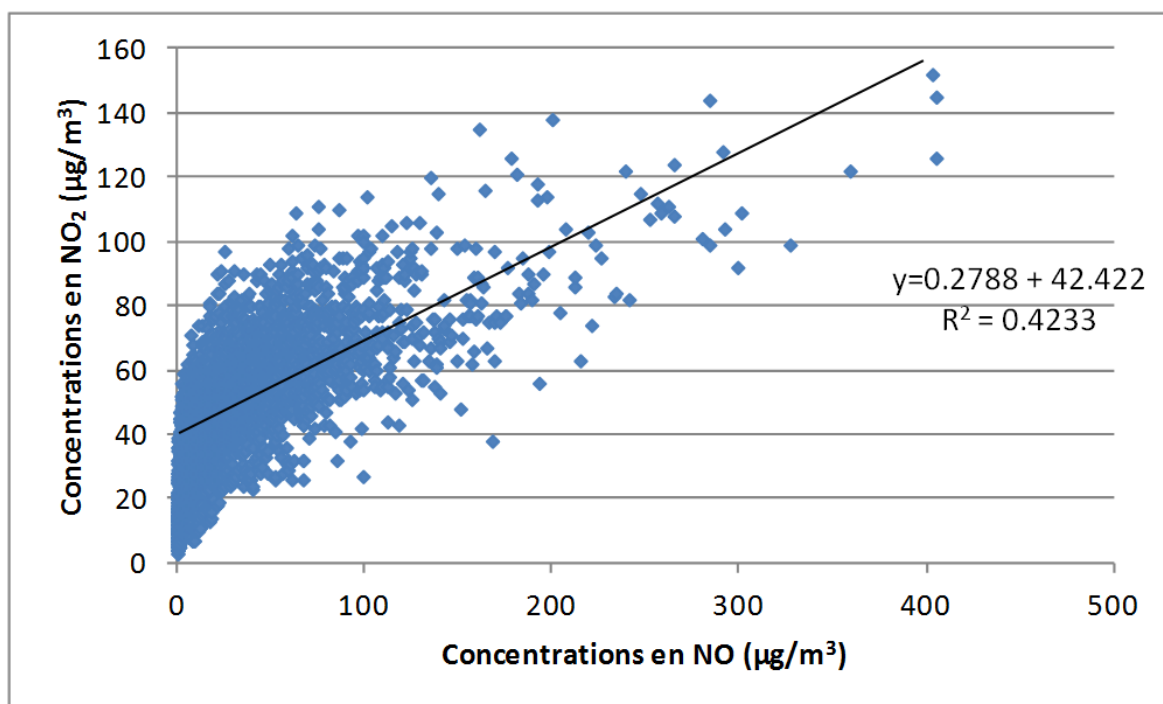
Station de Roubaix (trafic) :

- Relation entre NO_2 et NO , y et R^2 établies pour $\text{NO} > 20 \mu\text{g}/\text{m}^3$:



Station de Valenciennes (trafic) :

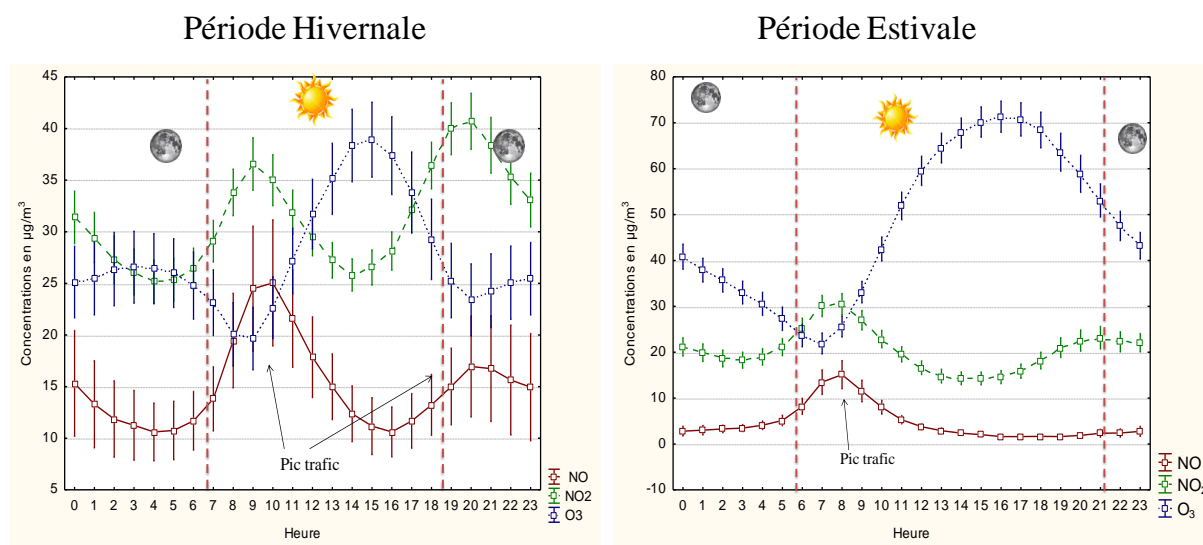
- Relation entre NO_2 et NO , y et R^2 établies pour $\text{NO} > 20 \mu\text{g}/\text{m}^3$:



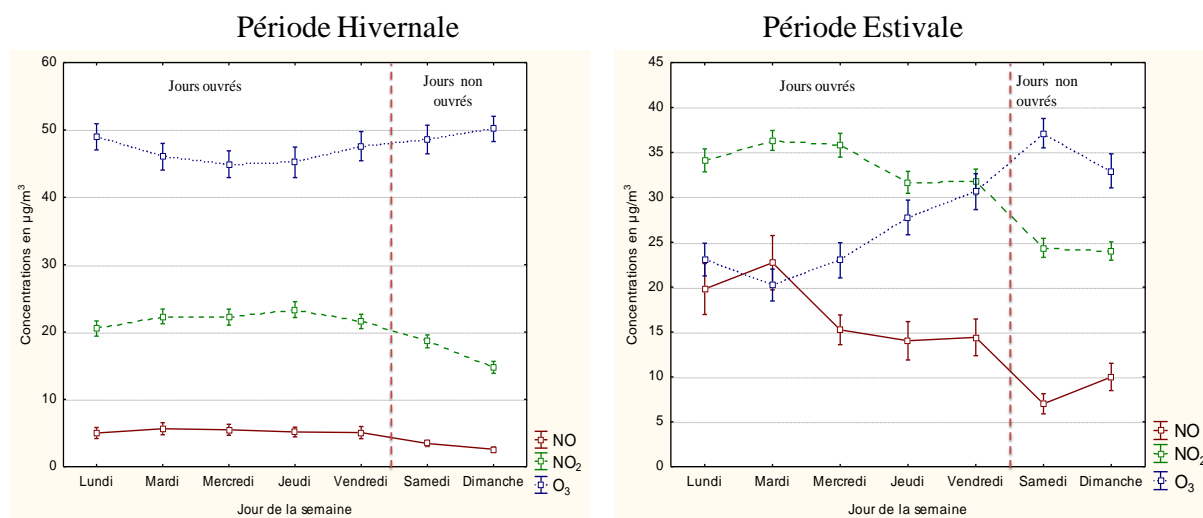
Annexe C : Forme journalière et hebdomadaire de pollution :

Site de Douai (Urbain) :

Forme journalière :



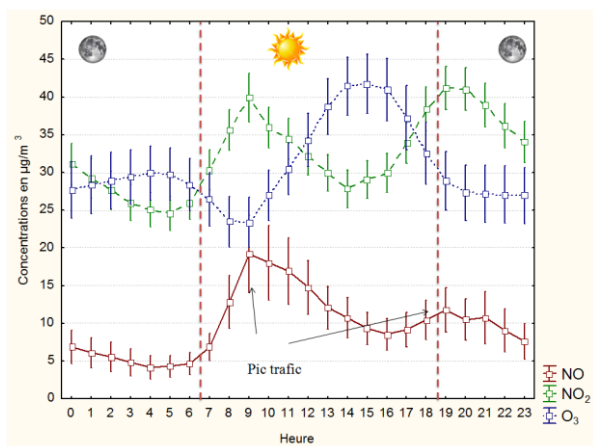
Forme hebdomadaire :



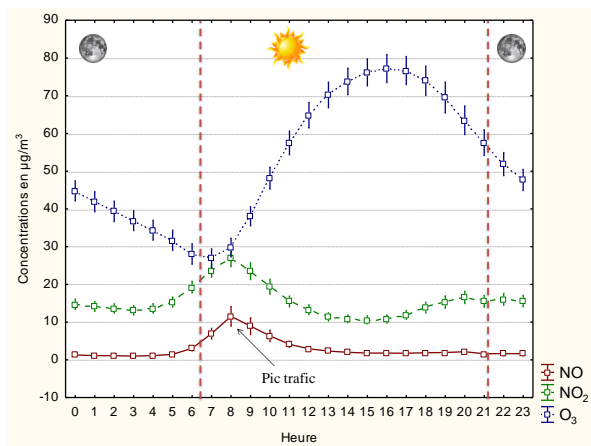
Site de Béthune (Urbain) :

Forme Journalière :

Période Hivernale

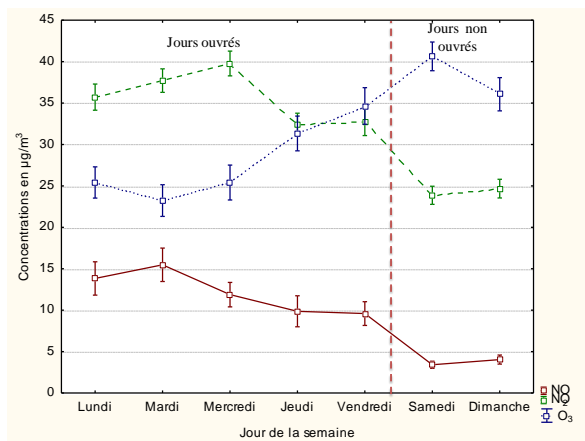


Période Estivale

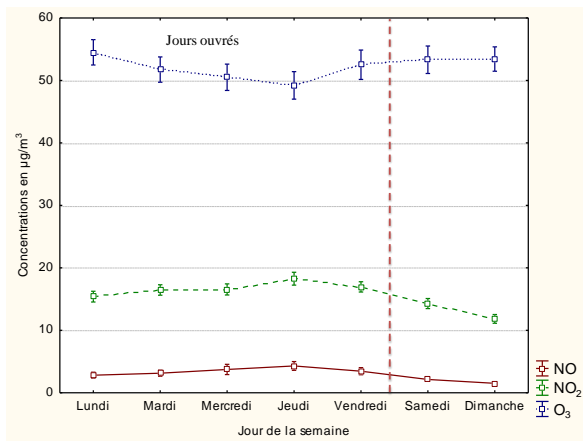


Forme hebdomadaire :

Période Hivernale

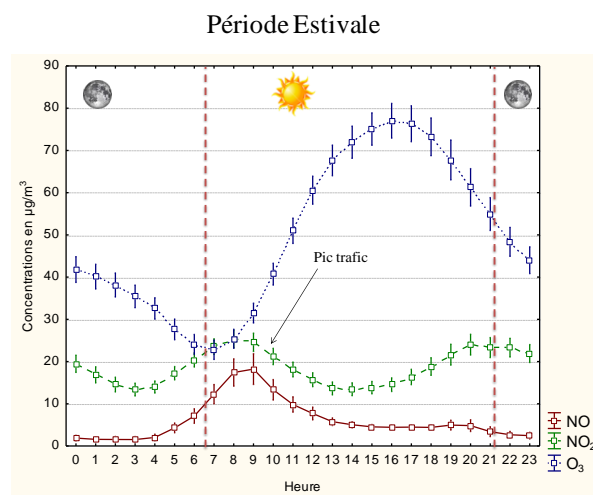
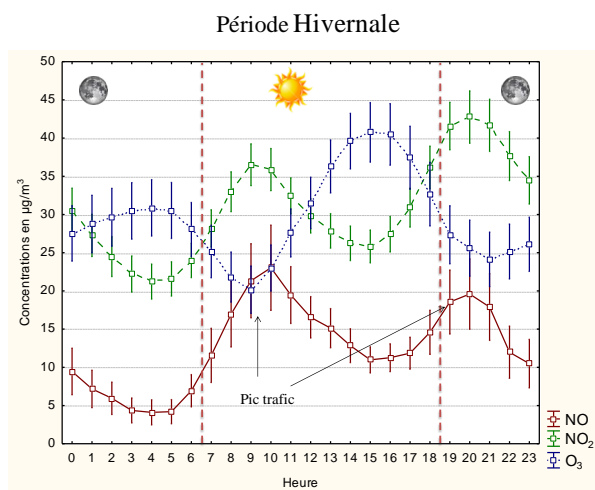


Période Estivale

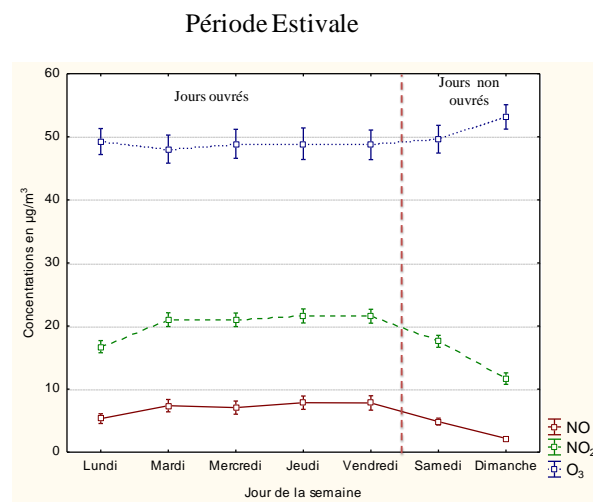
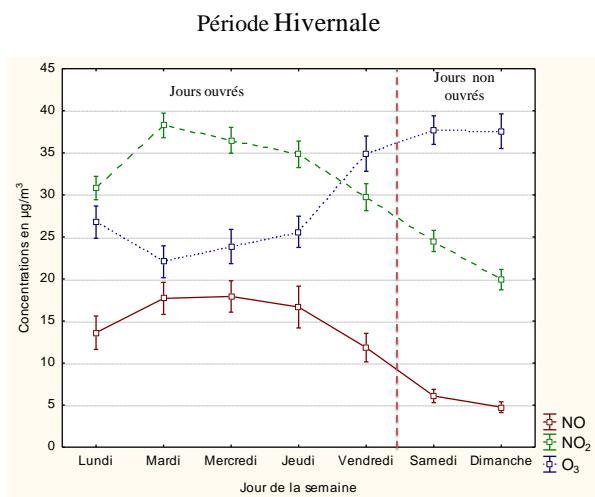


Site de Creil (Urbain) :

Forme journalière :

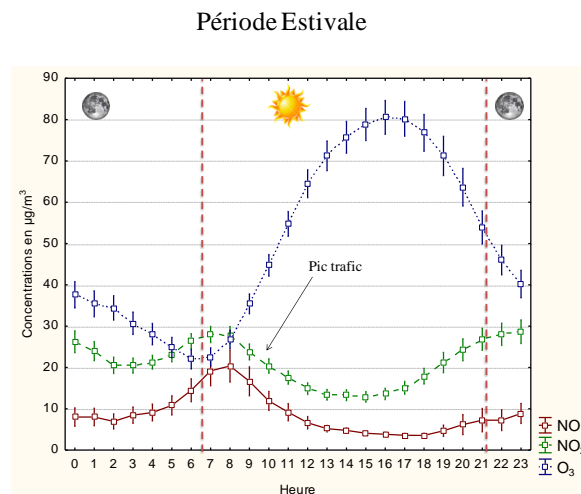
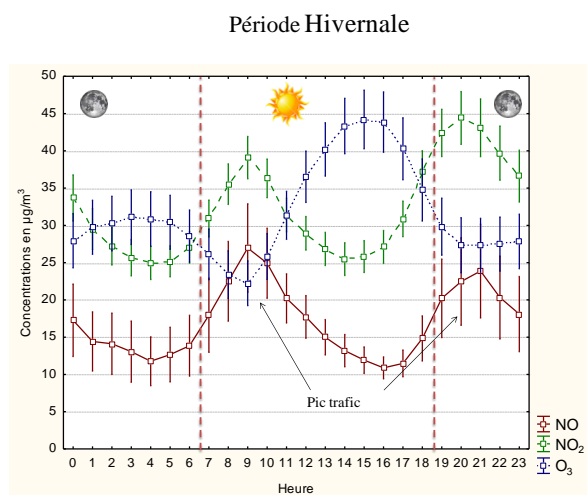


Forme hebdomadaire :

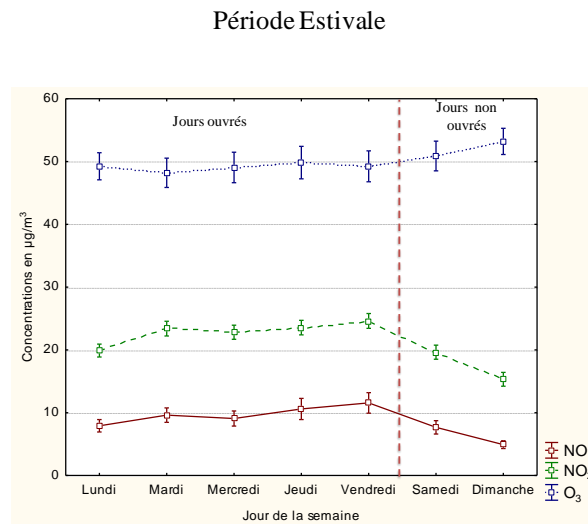
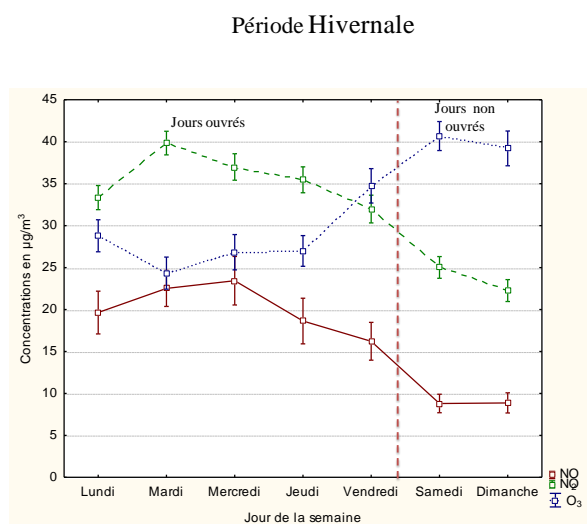


Site de Nogent sur Oise (Périurbain) :

Forme journalière :

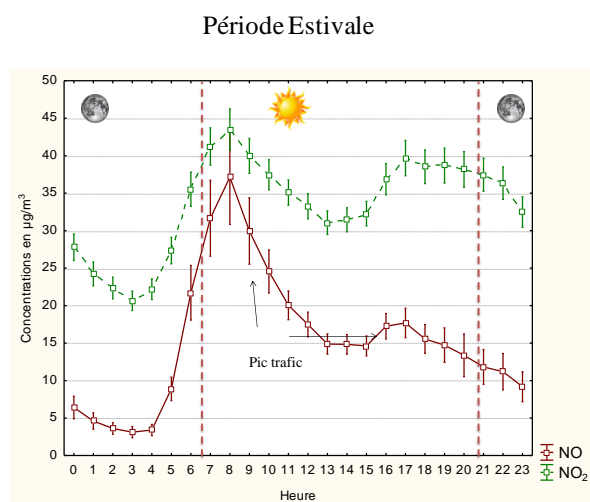
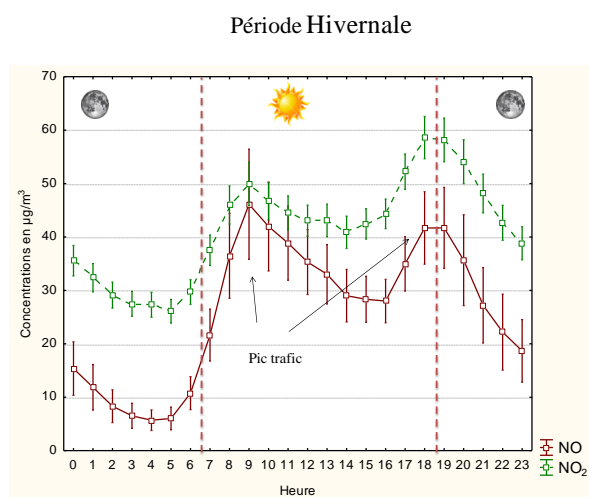


Forme hebdomadaire :

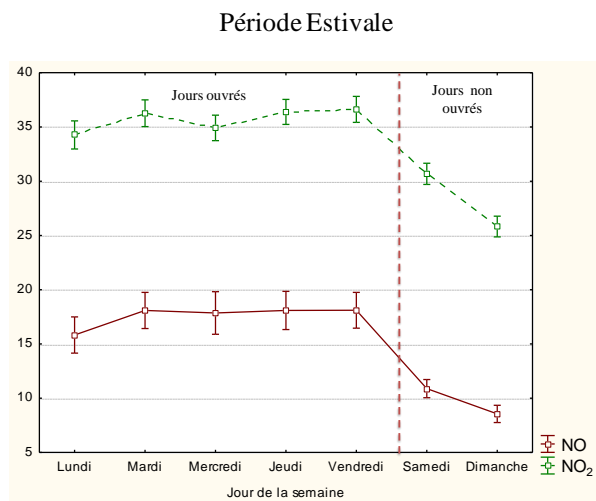
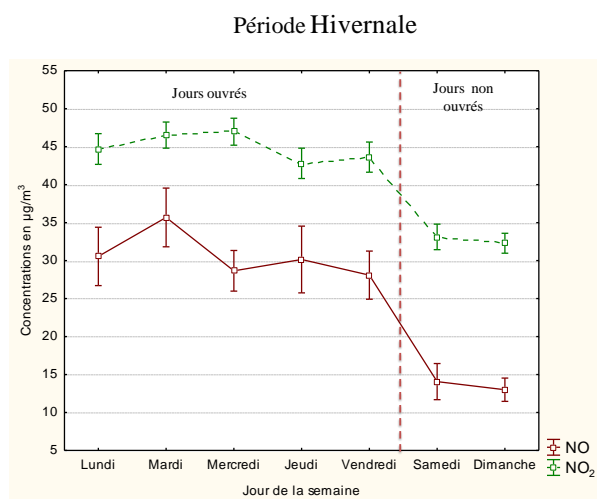


Site Valenciennes (Trafic) :

Forme journalière :

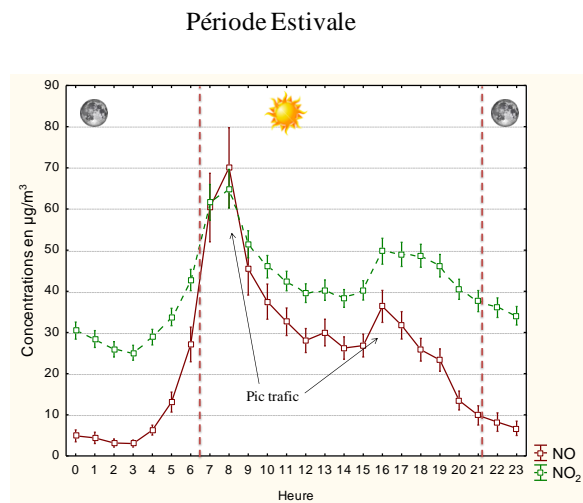
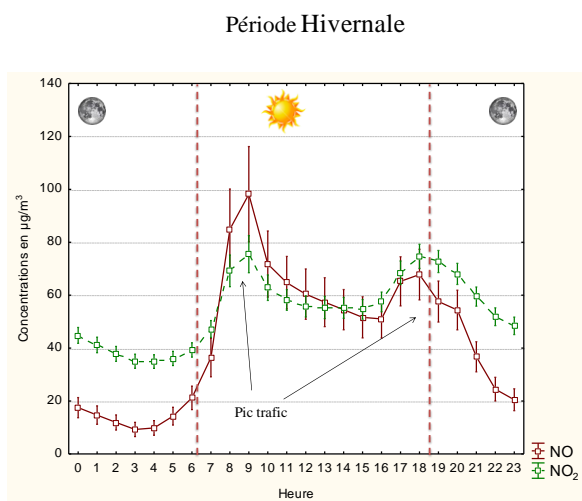


Forme Hebdomadaire :

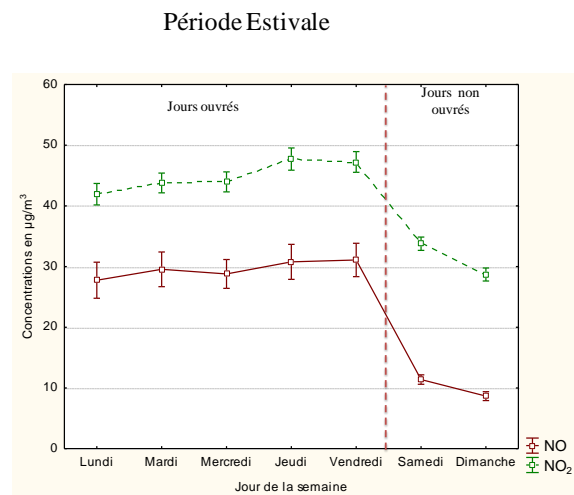
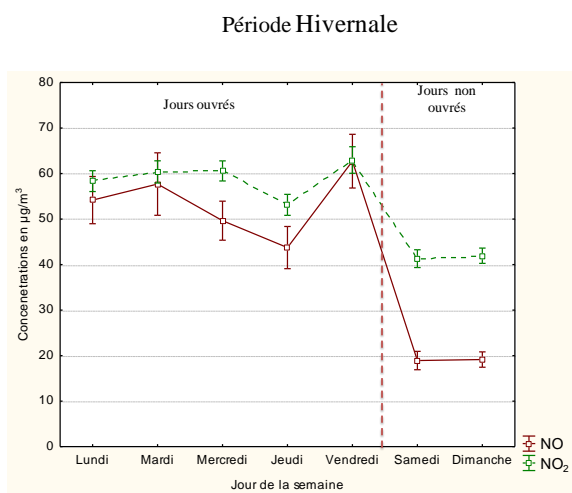


Site de Roubaix (Trafic) :

Forme journalière :

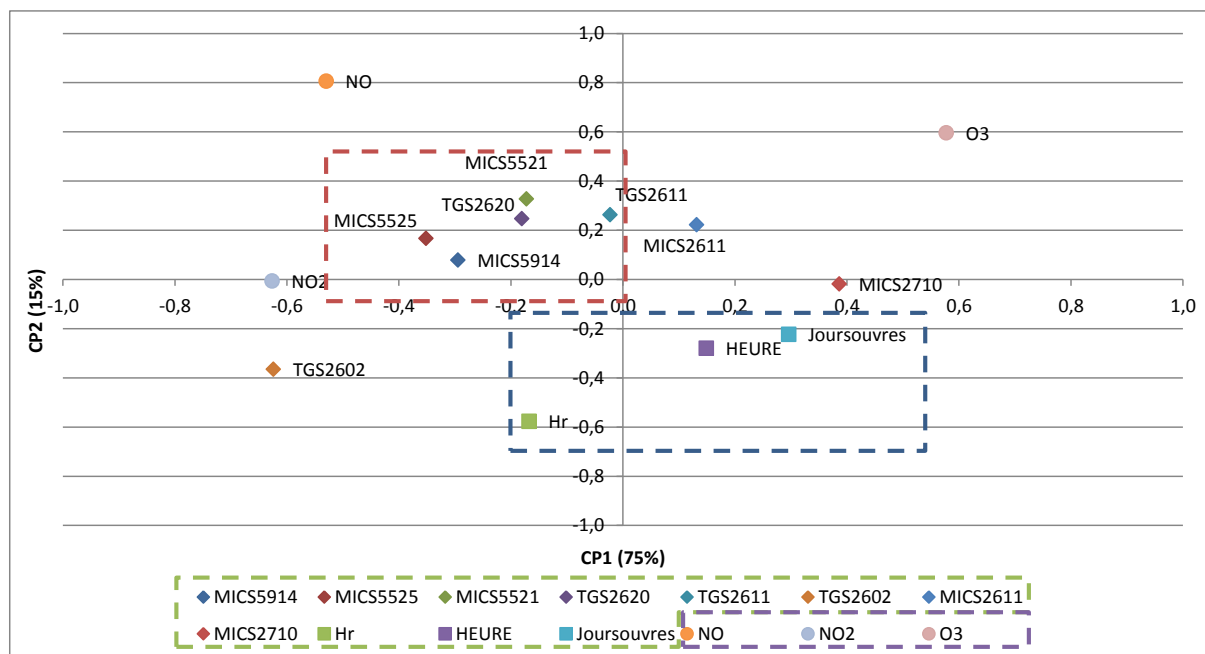


Forme hebdomadaire :



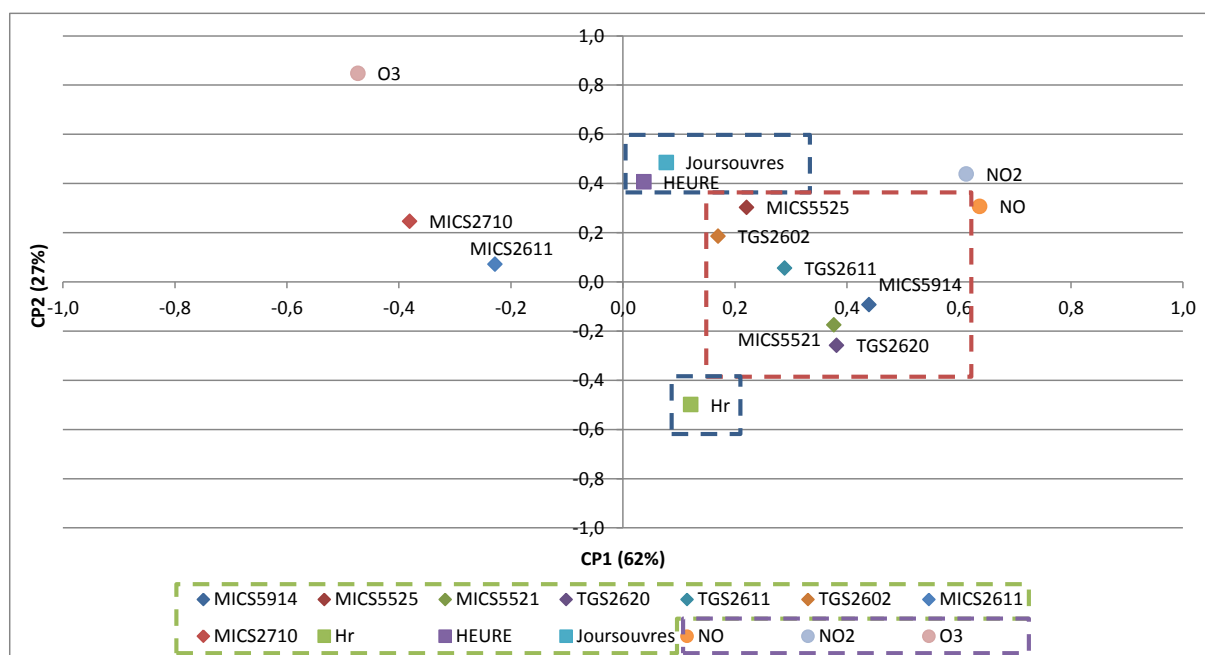
Annexe D : Etude des comportements des capteurs :

Site de Béthune (urbain) :



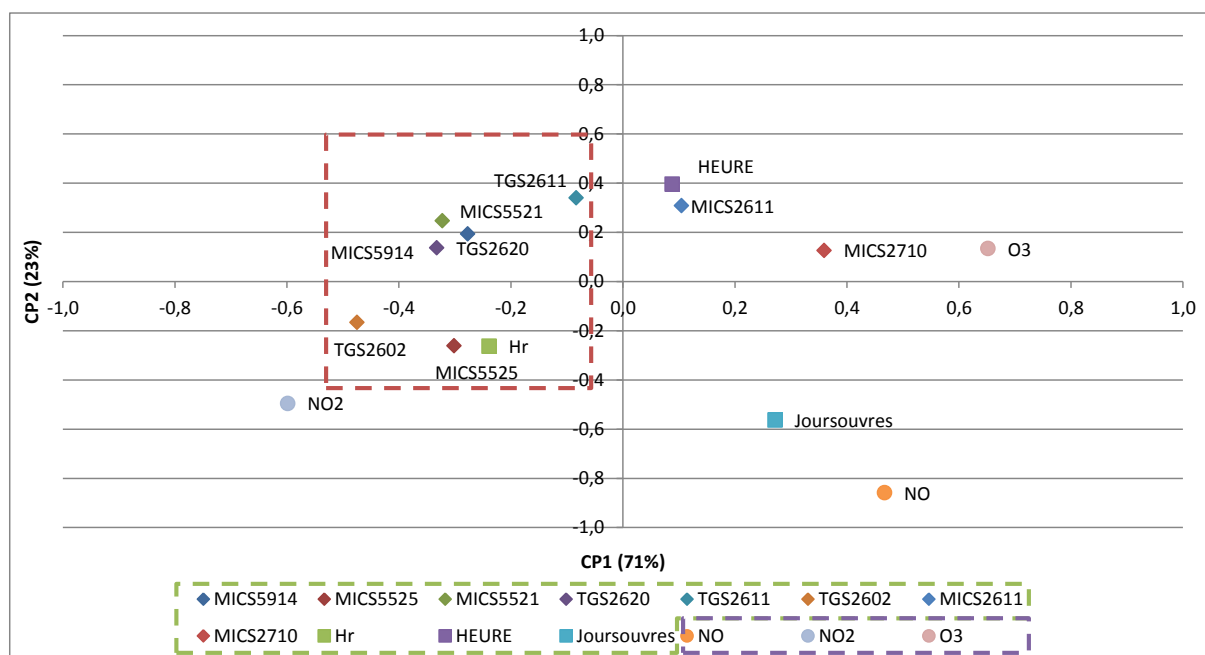
Dans cet exemple, nous remarquons des comportements similaires à ce qui a été présenté dans la partie 4.1. Le capteur TGS2600 a été éliminé à cause de la redondance de sa réponse avec le TGS2620. Nous observons :

- les oxydes d'azote (NO, NO₂) sont corrélés entre eux
- l'O₃ est anti-corrélé avec les oxydes d'azotes
- le capteur Mics2611 (capteur choisi pour la détection d'O₃) et le capteur Mics2710 (capteur choisi pour la détection de NO₂) ont des réponses corrélées avec l'O₃.
- Les capteurs Mics5521, TGS 2620, TGS 2602 et Mics5914, ont des réponses corrélées avec la présence des oxydes d'azotes. Ainsi ces capteurs sont corrélés avec des différents polluants réducteurs qui proviennent de la même source de NO et NO₂ (comme le CO et les Hydrocarbures).
- Les variables ajoutées (Hr, Heure, Jo) contribuent d'une façon plus faible au système. Par contre, nous pouvons les corrélérer avec la présence de l'O₃, c'est à dire la pollution photochimique présente durant la journée.

Site de Creil (Urbain) :

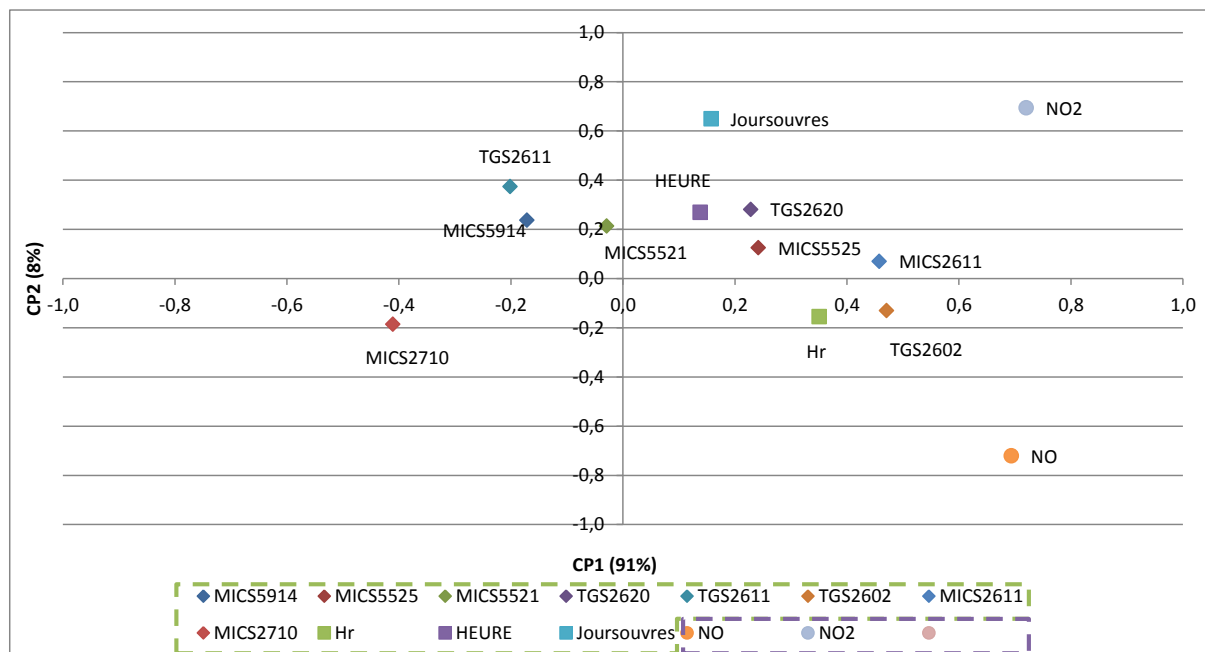
Dans cette étude, nous remarquons aussi des comportements similaires à ce qui a été présenté dans la partie 4.1. :

- les oxydes d'azote (NO, NO₂) sont très corrélés entre eux
- l'O₃ est anti-corrélé avec les oxydes d'azotes
- le capteur Mics2611 (capteur choisi pour la détection d'O₃) et le capteur Mics2710 (capteur choisi pour la détection de NO₂) ont des réponses corrélées avec l'O₃.
- Les capteurs TGS 2620, TGS 2602, TGS 2611, Mics5914, Mics5521, et Mics5525 ont des réponses corrélées avec la présence des oxydes d'azotes. Ainsi ces capteurs sont corrélés avec des différents polluants réducteurs qui proviennent de la même source de NO et NO₂ (comme le CO et les Hydrocarbures).
- L'humidité de l'air est indépendante des autres variables, ce qui indique la présence des pluies
- Les variables ajoutées (Heure, Jo) sont corrélés avec la présence de l'O₃, NO et NO₂.

Site de Douai (Urbain) :

Nous observons dans cet exemple :

- les oxydes d'azote (NO, NO₂) ne sont plus très corrélés entre eux, ce qui indique la présence des sources différentes pour NO et NO₂
- l'O₃ est anti-corrélé avec les oxydes d'azotes
- le capteur Mics2611 (capteur choisi pour la détection d'O₃) et le capteur Mics2710 (capteur choisi pour la détection de NO₂) ont des réponses corrélées avec l'O₃, ainsi que l'heure du jour.
- Les capteurs TGS 2620, TGS 2602, TGS 2611, Mics5914, Mics5521, et Mics5525 ont des réponses corrélées avec la présence de NO₂. Ainsi ces capteurs sont corrélés avec des différents polluants réducteurs qui proviennent de la même source NO₂, par contre ces capteurs ne sont plus corrélés à la source de NO (comme le CO et les Hydrocarbures).
- Les variables ajoutées (Heure, Jo, HR) Sont indépendants entre eux. Avec la variable des jours ouvrés est corrélée au NO.

Site de Valenciennes (Trafic) :

Ce site est un site trafic, d'où l'absence de l'O₃ d'où nous ne trouvons que les variables d'oxydes d'azotes. Nous avons ainsi obtenue deux composantes principales, la première possédant une variance de 91% et la deuxième possède une variance de 8%.

- les oxydes d'azote (NO, NO₂) sont très corrélés entre eux selon la première composante CP1 qui contient presque 91% des informations
- le capteur Mics2611 est très corrélé aux oxydes d'azotes (selon CP1), par contre Mics2710 est corrélé négativement à ces gaz (selon CP1)
- Les capteurs TGS 2620, TGS 2602, TGS 2611, Mics5914, Mics5521, et Mics5525 ont des réponses corrélées plus à NO₂ que NO.
- Les variables ajoutées (Heure, Jo, HR) ne présentent pas une corrélation claire.

NNT : 2012 EMSE 0677

Zaher AL BARAKEH

Suivi de pollution atmosphérique par système multi-capteurs – méthode mixte de classification et de détermination d'un indice de pollution.

Spécialité: Génie de Procédé

Mots clefs : Pollution atmosphérique, capteurs de gaz semi-conducteurs, réseaux de capteurs, logique floue.

Résumé :

Cette thèse a pour objectif le développement d'un système multi-capteurs de gaz permettant une évaluation en continu et en temps réel des différentes types de pollution atmosphérique en zone urbain, en classifiant notamment les pollutions de type urbaine, photochimique, ou encore liée au trafic. Le projet se base sur l'utilisation de différents capteurs de gaz de type semi-conducteur disponibles dans le commerce qui sont intégrés dans un dispositif autonome et portable, afin qu'il puisse fonctionner sur site.

Dans un premier temps, et en grande partie à Saint-Etienne, différents types de capteurs sont sélectionnés puis leurs performances sont testées sur un banc simulant les atmosphères polluées et développé pour l'occasion. Afin de pallier aux problèmes de non répétabilité et de dérive de la ligne de base et de la sensibilité, des procédures de prétraitement de standardisation sont mises au point.

Dans un deuxième temps, et en grande partie à Douai, différents sites de tests sont identifiés et leurs historiques de pollution sont étudiés. Plusieurs campagnes en stations de mesure d'une semaine, recouvrant les différentes saisons et les différents types de sites, sont alors menées. Il y est collecté conjointement les signaux des capteurs et des analyseurs de gaz réglementés. Des méthodes basées sur les réseaux de neurones sont alors appliquées afin d'obtenir conjointement, à partir des signaux des capteurs, une classification parmi 3 types de pollutions (urbaine, trafic et photochimique) ainsi qu'un indicateur global de qualité de l'air. Ces méthodes utilisent une approche basée sur la logique floue afin d'éviter les problèmes d'effet de bord.